THE PRODUCTION AND IDENTIFICATION

OF MANDARIN TONES IN CONTEXT

Meng Liu

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Master of Arts
in the Department East Asian Languages & Cultures
Indiana University
Spring 2015

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Master of Arts.

Master's Project Committee

Chien-Jer Charles Lin, Chair.

Natsuko Tsujimura

Isabelle Darcy

Copyright © 2015
Meng Liu

Acknowledgments

First, I want to express my immense gratitude for all of the academic support, draft revision, project coordination, and question answering from my Master's Thesis Committee, including Professor Chien-Jer Charles Lin (Committee Chair), Professor Natsuko Tsujimura, and Professor Isabelle Darcy. Moreover, I want to thank professor Jennifer Li-Chia Liu, who initially offered me the invaluable opportunity to study and teach at Indiana University, which gave me the foundation to advance my academic career here in the United States. Also, I want to thank Professor Chilin Shih, who helped me find the related research articles, and Professor Kenneth J. de Jong, and Professor Yasuko Watt, as well as Professor Wen Cao (Beijing Language and Culture University), who helped me shape this project.

Second, I want to extend my gratitude to Aaron Lee Albin, who offered critical technical and personal support, and who helped me with the experimental design, DMDX scripts, and R scripts. Also, to Stephanie Dickinson (Indiana Statistical Consulting Center), who lent her talents to helping me with analyzing the statistics.

Furthermore, I want to thank Lara for these past two years' help with answering numerous questions concerning degree requirements, coursework, TA-ships, and MA projects, and Professor Scott O'Bryan for helping me forming the committee members.

In addition, I want to thank my parents who have educated me on how to become a decent human being, nurtured my desire for knowledge, and persistently supported me. Also, I appreciate the everyday help from Jonathan Cox, who has greatly contributed to improving my academic English, refining my project writing, and taking care of me over these two years' of study at IU. Finally, thanks to those people who have been a part of my life, supported me, and helped me to move forward.

Meng Liu

THE PRODUCTION AND IDENTIFICATION OF MANDARIN TONES IN CONTEXT

The purpose of the thesis is to study the acoustic variation of Mandarin tones produced in context by native Mandarin speakers (Study I) and how these acoustic variations of Mandarin tones influence L2 learners' identification of Mandarin tones (Study II).

Study I revealed that both F0 contour and F0 height of the two tone sequence of disyllabic non words from L1 Chinese were influenced by both syllable position and tonal context (conflicting/compatible context).

Study II tested the identification of Mandarin tones by L1 English learners of Chinese in both monosyllable and two tone sequence of disyllabic non-words by using DMDX (a software for the experimental control and timing of stimulus display); in terms of how accuracy rates, identification sensitivity, error patterns, and reaction times are influenced by tonal context, syllable position, and learning experience. This study found that both syllable position and context affected the tone identification of L1 English learners of Chinese. Tones in monosyllables were identified with the highest accuracy and sensitivity, and shortest reaction time, followed by tones in the final syllable and tones in the initial syllable. Additionally, fewer errors were made in the compatible context than the conflicting context. With more learning experience, the effect of the compatible/conflicting context decreased for both tones in the initial syllable and final syllable tasks. The identification accuracy and sensitivity of Tone 1 (H) and Tone 4 (HL) were better than Tone 2 (LH) and Tone 3 (L) among the three tasks. The confusion between Tone 2 (LH) and Tone 3 (L) was most salient.

This thesis helps fill the current knowledge gap concerning L2 learners' identification difficulty of two tone sequence of Mandarin lexical tones, caused by the acoustic variation that existed in native Mandarin speakers' production. This new information contributes to a deeper understanding of context effect on Mandarin tone identification of L2 learners. This can benefit teachers in predicting the points of difficulty in learning Mandarin tones and assist students to improve their tone identification in context.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

1.1 Background: Mandarin Tones

In Mandarin Chinese, most morphemes are monosyllables. Traditionally, there exist about 400 syllable types with four lexical tones and lexically meaningful pitch patterns in Mandarin Chinese. The tones are numbered as Tone 1, Tone 2, Tone 3, and Tone 4.

Most of Mandarin Chinese words are disyllabic in contemporary usage (Duanmu, 1999). The study further suggests that due to the increase of the new vocabulary and homonym avoicdance, disyllabic words have increased drastically in the past 1000 years.

Researches have studied Mandarin Tones mainly from the following perspectives: the production and perception of Mandarin lexical tones by L1 Mandarin speakers, as well as the acquisition of Mandarin lexical tones by L1 English speakers. These aspects will be examined in Chapter 2.

1.2 Purpose of the Present Study

The purpose of this study is to explore the production and identification of Mandarin tones presented in monosyllables and tone sequence of two with a focus on the influence of tonal contexts. Specifically, this study tries to explore the tonal variations in context produced by L1 Chinese speakers and how L1 English learners of Chinese perceive these nonnative tones in context.

The results of this study are primarily beneficial to Chinese language learning and classroom teaching by identifying tonal combinations and contexts that pose problems for production and/or identification. Additionally, they can contribute in establishing the framework of assessing students' oral production and listening comprehension concerning tones. As a final practical application, they can be used in developing textbooks and training materials for learners of all levels and Chinese programs, thus allowing training programs to be developed in the aforementioned areas.

1.3 Organization of the Study

The rest of the thesis is organized into four subsequent chapters. Chapter 2 reviews the literature on Mandarin tone production and perception from both the L1 Chinese and L2 learners' perspectives. Chapter 3 presents the study of the production of mandarin tones in context by L1 Chinese speakers. Chapter 4 explores the identification of mandarin tones in context by L1 English learners of Chinese. Finally, a conclusion is offered in Chapter 5.

CHAPTER 2

LITERATURE REVIEW

This section provides a review of Mandarin tone production and perception from two aspects. One is the native tone production and perception by L1 Mandarin Chinese speakers and the focus is the context effects of native Mandarin tone production. The other is the nonnative tone production and perception by L2 learners and the focus is how nonnative speakers perceive Mandarin lexical tones in context.

2.1 The Production and Perception of Mandarin Lexical Tones by L1 Mandarin Speakers

In this section both Mandarin lexical tone production and perception by L1 Mandarin speakers are discussed. The production and perception of Mandarin lexical tone in isolation as well as in context are discussed.

2.1.1 The Production of Mandarin Lexical Tones by L1 Mandarin Speakers

When produced in isolation, Mandarin lexical tones seem well defined and quite stable, including F0 height and F0 contour (Xu, 1997). As shown in Figure 1 and Table 1, the four lexical tones can be described as Tone 1: High (55), Tone 2: Rising (35), Tone 3: Low (21) and Tone 4: Falling (51), by using the 5-level scale and tone contour can be expressed by sequence of tone letters indicating starting and ending points (Chao, 1930). The four lexical tones are alternately described as: Tone 1 (high-level), Tone 2 (mid-rising), Tone 3 (mid-falling-rising) and Tone 4 (high-falling) based on F0 height and F0 contour (Howie, 1976).

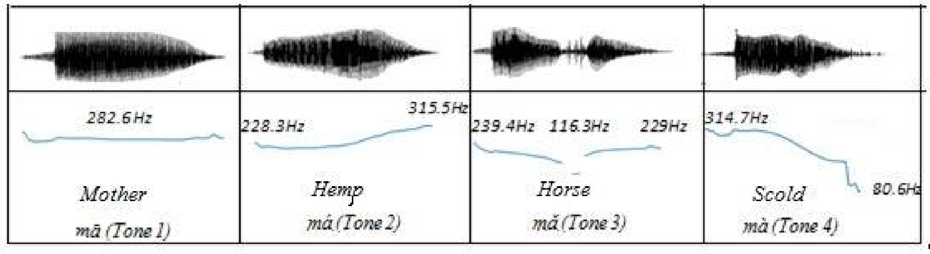


Figure1: Mandarin tones, recorded by a female native speaker from Beijing.

Table 1: Mandarin tones

	5-level scale(Chao, 1930)	F0 height and F0 contour(Howie, 1976)
Tone1	High (55)	High-level
Tone2	Rising (35)	Mid-rising
Tone3	Low (21)	Mid-falling-rising
Tone4	Falling (51)	High-falling

Through the description, we can see that F0 contour and F0 height are the primary difference among the four Mandarin lexical tones. Besides the difference in F0 contour and F0 height, these tones also differ in duration. Howie (1976) suggests that Mandarin tones appear to have some intrinsic durational differences: Tone 3 (L) tends to be the longest, and Tone 4 (HL) is the shortest. Tone 2 (LH) is generally shorter than Tone 3 (L), but longer than Tone 1 (H).

When produced in context, tone sequences produced by native speakers are highly influenced by the tonal contexts (Shih, 1987; Xu, 1997), though previous studies found the four underlying tones in connected speech of Mandarin Chinese are Tone 1 (H), Tone 2 (LH), Tone 3 (L), and Tone 4 (HL), according to their onset and offset F0 height (Shih, 1987; Shih, 1988; Shih & Sproat, 1992; Duanmu, 2000) (See table 2).

Table 2: Onset and offset values of Mandarin lexical tones

Onset offset	High	Low
High	Tone 1	Tone 4
Low	Tone 2	Tone 3

The variations caused by the surrounding tonal environment changing F0 height and F0 contour can be categorized into two categories, including phonological variants and phonetic variants.

Concerning the phonological variants, Chao (1968) posits that Mandarin lexical tones are mainly influenced by the following tones as indicated by the tone sandhi rules (phonological tonal changes based on adjacent tonal environment), including Tone 3 (L) sandhi rule, Tone 4 (HL) sandhi rule and Tone 2 (LH) sandhi rule. Tone 3 (L) sandhi rule is the most important sandhi rule in Mandarin Chinese. Specifically, there are two variants of Tone 3: a mid-rising half third tone (LH) and low-falling half third tone (L). The mid-rising half third tone (LH) is pronounced when it precedes another syllable carrying Tone 3 (L), whereas the low-falling half third tone (L) is pronounced when it precedes another syllable not carrying Tone 3 (L) (i.e. Tone 1, Tone 2, or Tone 4). In addition, Tone 4 (HL) becomes high mid tone when followed by another Tone 4 (HL). Tone 2 (LH) becomes Tone 1 (H) when preceded by Tone 1 (H) or Tone 2 (LH) and followed by any other stressed tone .

Many studies have shown that Tone 3 in continuous speech is realized as low and the final rise seen in the citation form is usually absent in non-prepausal positions (Shih, 1988, 1992; Xu, 1993,1994, 1997; Duanmu, 2000, 2007; Hallé et al., 2004). Duanmu (2000, 2007) suggests that Tone 3 (L) is a low-falling tone in nonfinal position and a low-falling or low-dipping tone in utterance-final position. Hallé et al. (2004) confirmed that Tone 3 (L) is most often pronounced as one of these "half third tones" in the non-prepausal position in continuous speech. It is important to notice that the surface form of mid-rising half third tone (LH) is very similar to Tone 2 (LH), though difference exists. Peng (1996, 2000) explored Mandarin Tone 3 sandhi by using the two syllables of tone sequences, composed with Tone 2 (LH) and Tone 3 (L) versus

Tone 3 (half third tone) and Tone 3 (L). It is found that overall F0 of the derived mid-rising half third tone from Tone 3 is slightly lower than the underlying Tone 2 (LH), although the tone shape of the sandhi tone is the same as that of the underlying Tone 2 (LH). In addition, the surface form of low-falling half third tone is similar with Tone 4 (HL), though difference exists. Gårding et al. (1986) explored native Mandarin speakers' production of Tone 3 (L) and Tone 4 (HL) in connected speech and found that Tone 3 (L) is connected with a low pitch level throughout the second half of the vowel and Tone 4 (HL) is connected with a gradual fall over the main part of the vocalic segment.

Concerning the phonetic variants, it has been reported that both the F0 height and F0 contour are influenced by the tonal context in Mandarin (Shih, 1987, 1988; Shen, 1990, Xu, 1993, 1994, 1997). Shih (1987), through examining tone sequences in Mandarin, found that the initial target of a tone serves as a transitional region between the preceding tone and the following tone. Shih (1988) further described that when different tones are produced together in words and sentences, the situation often arises that adjacent tonal targets have opposite values that is where tonal co-articulation is expected. That is to say the phonetic variants are the variability caused by tonal co-articulation when Mandarin lexical tones vary in their acoustic realizations depending on the tones of the preceding and following syllables.

Previous researches offered a variety of views of how F0 height and F0 contours of Mandarin lexical tones are influenced by the surrouding tonal context, specifically concerning whether the context effect is symmetric or asymmetric. Shen (1990) suggested that tonal coarticulation in Mandarin were symmetric bi-directional effects, as the tonal contours undergo certain variations depending on preceding and following tones,by using Mandarin tri-tonal combinations. By using more controled experimental design, Xu (1997) found that acoustic variations of Mandarin tones are due to two assymetric effects: carry-over effects and anticipatory effects by using tone sequences of two sylables with and without carrier phrases. The study described the carry-over effects as the onset of a tone is assimilated into the offset value of the previous tone and the anticipatory effects as a low onset value of a tone raises the maximum F0 value of a preceding tone. These two effects differ both in nature and in magnitude. It is suggested that the nature of carry-over effects are assimilation and the nature of anticipatory effects are disassimilation. The magnitude of carry-over effects are larger and the the magnitude of anticipatory effects are smaller. That is to say the contour of a tone is significantly influenced by the offset of the preceding tone but not significantly influenced by the onset of the following tone. Xu (1993, 1994) explored trisyllabic tone sequences in Mandarin and divided them into compatible and conflicting contexts. A compatible tone sequence, compatible context, means that the end of the preceding tone and the beginning of the following tone match in pitch height. A conflicting tone sequence, conflicting context, means that the end of the preceding tone and the beginning of the following tone do not match in pitch height. The study found that the amount of deviation of a tone from its canoical form due to coarticulation varies depending on the nature of the tonal context. The deviation is relatively smaller in compatible context and greater in conflicting context. Therefore, we can see that tonal context is the main force causing the variablity in the surface form of the tonal sequences. When analyzing production of tonal sequences, the carry-over effects and anticipatory effects, as well as compatible context and conflicting context should be included. This combined view of tonal co-articulation concerning tone sequence of disyllabic words in Mandarin has not been systematically examined in the previous studies.

In addition, in natural speech, studies have explored how syllable position and stress would influence the surface form of utterances produced in context. Shih & Sproat (1992) demonstrated that a syllable's ability to resist tonal coarticulation from adjacent tones is related to its prosodic strength. Tones on prosodically weak syllables tend to have less extreme tone shapes, or, greater undershoot. Lin et al. (1984) explored the stress pattern of normal disyllabic words in Mandarin and discovered that among most isolated disyllabic words, the final syllable is more heavily stressed than the initial ones. Wang (2004) discovered that when appearing in the final position in a disyllabic word, Tone 1 (H), makes a syllable have the smallest possibility to be unstressed at lexical level, and also makes it most prominent in perception and the Tone 3 (L), does the reverse. Based on the aforementioned studies, Tone 1 (H) in the final position of the disyllabic words should be with the highest ability to resist tonal co-articulation and Tone 3 (L) in the initial position of the disyllabic words should be with the lowest ability to resist tonal co-articulation.

Besides aforementioned the acoustic analysis of the tonal production in contexts, several studies (Xu & Wang, 2001, 2003; Xu & Sun, 2002; Xu, 2004; Xu, 2006) are trying to explore the reason of the substantial variability of the surface form of the tones from the articulator factors. It is suggested that there are two articulatory constraints, including maximum speed of pitch change and coordination of laryngeal and supralaryngeal movements. Xu & Wang (2003) suggested that due to these constraints, the implementation of a simple pitch target may result in surface F0 forms that only partially reflect the underlying pitch targets. The study proposed a preliminary framework for accounting for certain surface F0 variations in speech: Target Approximation model, which consists of pitch targets and rules of their implementation, by taking into consideration the articulatory constraints on the production of surface F0 contours.

Xu (2006) summarized that the target of a tone is compromised due to articulatory easiness, the changed underlying tonal target, as well as other linguistic functions, including articulatory strength and syllable duration, etc. Xu (2004) suggested that the model is based on the new understanding of the interaction between communicative demands of the speech and articulatory constraints on speech production. The interaction between the two factors could form robust variations in the F0 contours of tones. According to Xu (2004), the core of the model is the assumption that phonological tone categories are not directly mapped onto surface phonetic patterns. Instead, each tone is associated with an ideal pitch target that is physically possible to produce, as illustrated in the following figure (See Figure 2), a schematic sketch of the Target Approximation model. From the figure we can see that the phonological tone targets low and high tones, indicated by the dashed lines, are not directly implemented. The thick curve represents the F0 contour that results from articulatory implementation of the pitch targets, reflecting the articulatory constraints of tone production. This model can be employed to explain a series of tonal changes, including tone sandhi and tonal co-articulation, which refers to tonal variants that are strictly conditioned by tonal contexts and are phonetically motivated.

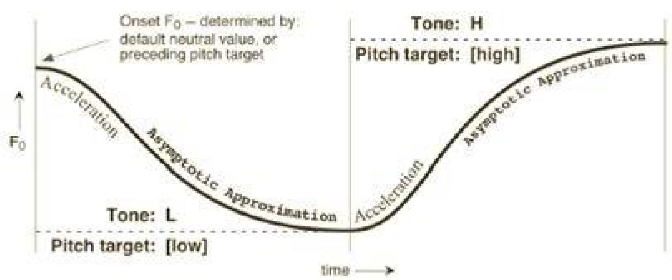


Figure 2: A schematic sketch of the Target Approximation Model (Xu, 2004).

Learning experience is another factor that influences L1 Mandarin speakers' production of Mandarin lexical tones both in isolation as well as in context. In general, L1 Mandarin-speaking children's tone production is not adult like. The acquisition of tone order and the learning difficulty is being discussed here. Li & Thompson (1977) did a longitudinal study in Taipei with 17 children from Mandarin-speaking families. The study found that the Mandarin high-level and falling tones are acquired before the rising and dipping tones, the rising and dipping tones are substituted for each other throughout the tone accuqisition process by studying children of Mandarin-speaking families. Concerning learning difficulty, Tone 2 (LH) and Tone 3 (L) are more difficult than Tone 1 (H) and Tone 4 (HL) for L1 Mandarin speaking children in general both in isolation and in contexts, and tonal contexts do play in a role of the children's production accuarcy (Li & Thompson, 1977; Wong et al., 2005; Wong, 2008; Wong, 2012a; Wong, 2012b). Not only do Tone 2 (LH) and Tone 3 (L) appear later in speech, but they are also the only tones confused with each other throughout much of the early acquisition period (Li & Thompson, 1977). Wong et al. (2005) and Wong (2012a) compared monosyllabic Mandarin lexical tones produced by 3-year-old Mandarin-speaking children growing up in Taiwan and in the United States. The study found that none of the four tones produced by the Mandarin-speaking children growing up in the United States and in Taiwanwas are adultlike. Overall, the production accuracy Tone 4 (HL) and Tone 1 (H) is higher than Tone 2 (LH) and Tone 3 (L). And the problem tone pairs include Tone 1 (H) and Tone 2 (LH), Tone 2 (LH) and Tone 3 (L), as well as Tone 4 (HL) and Tone 3 (L). Wong (2012b) study explored acoustic characteristics of three-year-olds' correct and incorrect monosyllabic Mandarin lexical tone productions and found that compared with adults' production, children's Tone 1 (H) is not as high or as level, and they have more difficulties producing low frequencies. Wong (2008) examined disyllabic tones produced by both Mandarin-speaking adults and 5- and 6-year-old Mandarin-speaking children that grew up in the United States and found that children's accuracy, different from adult, of the same tone varied depending on context. In addition, the study suggested that two tone combinations with more complex F0 contours are more difficult for children to produce. For example, the accuracy of the tone sequences of Tone 1 (H) and Tone 2 (LH) combination is much lower than the Tone 2 (LH) and Tone 1 (H) combination. This is highly possible due to the compatible and conflicting contexts, as discussed in Xu (1993, 1994).

Other than behavioral studies, neurophysiological studies have beed used to study the neural mechanisms of processing Mandarin tones. Previous research indicates that the left hemisphere is more adept at phonemic processing and while the right hemisphere is better at melodic and prosodic processing, including music, pitch contours, and affective prosody (as cited in Jongman et al., 2011). Liu et al. (2006) explored the production of Chinese lexical tones by native Chinese speakers by using fMRI (Functional magnetic resonance imaging) and their brain asymmetry analysis showed that tone production is left-hemisphere dominant.

To summarize, F0 height and F0 contour are the primary difference among the four lexical tones in Mandarin, and there are also duration differences. Tonal context, syllable position, and stress contribute to the variability of the actual tone production in the surface form. Also, L1 Mandarin-speaking children's tone production is not adult like. In addition, tone production is left-hemisphere dominant for L1 Mandarin speakers.

2.1.2 The perception of Mandarin Lexical Tones by L1 Mandarin Speakers

Many previous studies explored the perceptual cues, confusion patterns, and context effects in perceiving Mandarin lexical tones by L1 Mandarin speakers.

Concerning perceptual cues, it is found that both F0 height and F0 contour are essential dimensions of the perceptual cues to identify Mandarin tones by native Mandarin speakers (Massaro et al., 1985; Fu & Zeng, 2000; Kuo et al., 2008), while duration, amplitude and creaky voice are the secondary perceptual cues (Massaro et al., 1985; Tseng, 1990; Whalen & Xu, 1992;

Zeng, 2000; Kuo et al., 2008; Gottfried & Suiter, 1997). Massaro et al. (1985) found that both F0 height and F0 contour are the two essential dimensions of the perceptual cues, although the various acoustic cues are functionally integrated when Mandarin speakers identify the tones. Tseng (1990) stated that duration played a secondary role in tone perception, based on the acoustic analysis of Mandarin tones in isolation. Whalen & Xu (1992) examined the contribution of duration and amplitude contour to native Mandarin speakers' tone perception. It is showed that subjects were able to identify all but Tone 1 (H) tokens from the amplitude contours alone. Fu & Zeng (2000) explored three temporal envelop cues, including duration, amplitude countour and F0 contour. It was found that F0 is the primary acoustic cue and durational cue contributes mostly to discrimination of Tone 3 (L). Longer duration contributes to the identification of Tone 3 (L). The amplitude cue contributed to discrimination of Tone 3 (L) and Tone 4 (HL). Kuo et al. (2008) reconfirmed that F0 is the primary cue to perceive tonal contrast in Mandarin by Taiwanese Mandarin speakers, and the temporal coding of F0 and amplitude envelope both contributed somewhat to tone recognition, while duration had only a marginal effect. Also, native Mandarin-speaking listeners may be able to use acoustic information other than F0, such as the creaky voice quality in Tone 3 (L), to identify tones (Gottfried & Suiter, 1997). Creaky voice is a concomitant but not a necessary feature of Tone 3 (L) (Gårding et al., 1986).

Previous studies also explored the tone identification ability when the F0 contour is incomplete, by using limited acoustic input, including acoustically modified Mandarin tones (Lee et al., 2006, 2008) and gating techniques (Lee et al., 2001; Lai & Zhang, 2008). Lee et al. (2006, 2008) tested native Mandarin-speaking listeners' tone identification abilities by using acoustically modified Mandarin tones to explore the effects of limited acoustic input, including intact, silence-center, center-only and onset-only, and the role of the tonal context in tone identification. They found that the identification of onset-only syllables in isolation is the lowest among the examined conditions. Lee et al. (2006) found Tone 2 (LH) is identified least accurately. Listeners also took more time to identify tone 1 (H) and tone 2 (LH) than the other two tones. Lee et al. (2008) found that there is predominant confusion between Tone 2 (LH) and Tone 3 (L), and Tone 1 (H) and Tone 4 (HL). Lee (2001) explored the online perceptual processing of Mandarin tone by using a gating technique which provides information regarding the amount of sensory information required for the identification of tones by presenting fragments of progressively increasing duration (20ms). It is found that Mandarin listeners could correctly recognize a Mandarin tone well before the entire F0 contour of the tones was heard. Lai & Zhang (2008) employed a gating paradigm to explore the native Mandarin speakers' perception of Mandarin lexical tones. It is found that Tone 1 (H) has a significantly earlier Isolation Point (IP) than Tone 4 (HL), which has an earlier IP than Tones 2 (LH) and Tone 3 (L). Also, the result indicates that a hierarchy of cues at the onset of tonal identification was also found: F0 high > contour > F0 low. Specifically, high-onset tones, regardless of contours, are not misidentified as low-onset tones; but low-onset tones are sometimes misidentified as high-onset tones due to their contour shapes.

Many studies specifically explored the perceptual cues of native Mandarin-speaking listener's perception of Tone 2 (LH) and Tone 3 (L), as many previous studies found that Tone 2 (LH) and Tone 3(L) are mostly confused tone pairs for Mandarin-speaking listeners in both isolation forms and tone sequence of two (Whalen & Xu, 1992; Wang & Li, 1967; Peng, 1996). Whalen & Xu (1992) found that Tone 2 (LH) and Tone 3 (L) are mostly confused when only a brief segment is available to perceive. Wang & Li (1967) showed that Mandarin listeners cannot distinguish words and phrases with Tone 3 (L) and Tone 3 (L) sequence with Tone 2 (LH) and

Tone 3 (L) sequence. Peng (1996) confirmed that the derived Tone 2 (LH) from Tone 3 (L) and the underlying Tone 2 (LH) were perceptually indistinguishable to Mandarin speakers.

It is found that the perceptual cues for native Mandarin-speaking listeners' to distinguish Tone 2 (LH) and Tone 3 (L) includes, F0 height (Whalen & Xu, 1992; Shen et al., 2013), F0 contour: F0 turning point (inflection point of the tone), △F0 (the F0 difference between the tonal onset and the turning point) (Shen & Lin, 1991; Moore & Jongman, 1997), and duration (Blicher et al., 1990). Whalen & Xu (1992) found that low unchanging F0 is perceived as Tone 3 (L), indicating a partial effect of register in Mandarin. Shen et al. (2013) used an eye-tracking paradigm to examine time-sensitive perceptual processing of Mandarin Tone 2 (LH) and tone 3 (L). Native speakers of Mandarin listened to manipulated tone tokens and selected the corresponding word from four visually presented words. The study demonstrates the importance of pitch height at tone offset and turning point in the process of tone identification by native speakers of Mandarin. Tokens with high offset pitch were identified as Tone 2 (LH), and low offset pitch as Tone 3 (H). A low turning point pitch served as a pivotal cue for Tone 3 (L). The findings support the view that lexical tone perception is an incremental process, in which F0 height at critical points serves as an important cue. Shen & Lin (1991) and Moore & Jongman (1997) showed that two acoustic dimension of F0 turning point and the △F0 are important perceptual cues to differentiate Tone 2 (LH) and Tone 3 (L) for native Mandarin-speaking listeners. Shen & Lin (1991) found that later F0 turning point and larger △F0 contribute to perceive the continuum as Tone 3 (L). Blicher et al. (1990) found that the identification of Tone 2 (LH) decreased by elongating the syllable while keeping the proportion between the falling phase and the syllable duration. This indicates that the absolute duration of the falling phase is a durational perceptual cue for Mandarin-speaking listeners to distinguish Tone 2 (LH) and Tone 3 (L).

Previous studies (Lin & Wang, 1985; Gårding et al., 1986; Xu, 1993, 1994; Cao, 2010; Huang & Holt, 2008) also explored how the F0 height of the preceding and following tone and syllable position influences tone perception of native Mandarin-speaking listeners. When the initial syllable is the perception target, Lin & Wang (1985) found that tones are perceived relative to other tones. The study examined the context effect of perception Tone 1 (H) in Mandarin held at a consonant 115Hz, followed by Tone 4 (HL) varied onset F0 from 110 to 140 Hz. The study found that as the onset F0 in the second syllable increased, identification of the first Tone 1 (H) syllable as Tone 2 (LH) increased. This indicates that the offset of the target tone assimilate to the onset of the following syllable tone perceptually. When the following syllable tone is the perception target, it is found that the preceding tone functions as a reference tone and the effect is contrastive (Gårding et al., 1986; Huang & Holt, 2008). Gårding et al. (1986) suggested that the preceding tone, functioning as a clear reference was important for perceiving both Tone 3 (L) and Tone 4 (HL). The identification of Tone 4 (HL) was favored by an introductory rising or level part, and for Tone 3 (L) an introductory fall seemed to be important. Huang & Holt (2008) tested the effect of context on contour lexical tone perception, by examining Mandarin listeners' perception of Mandarin Tone 1 (H) and Tone 2 (LH). Results indicate that the mean F0 of a preceding sentence affects perception of contour lexical tones and the effect is contrastive. Following a sentence with a higher-frequency mean F0, the target syllable is more likely to be perceived as a lower frequency lexical tone and vice versa. Cao (2010) found that the recognition of a level tone is affected by both the preceding and the following tones. The study tested native Mandarin speakers' Mandarin tone perception of Tone 1

(H, /55/) in isolation as well as in contexts by using level tones with different pitch range/11/, /22/, /33/, /44/, /55/. With a preceding tone, even the lowest level contour is recognized as Tone 1. However,when it precedes the other tones, the lowest level contours would be recognized as "Half Tone 3", Tone 3 (L). Xu (1993, 1994) further examined the tonal context effect formed by the preceding and following tones as a whole, including compatible contexts and conflicting contexts. The study examined the L1 Chinese speakers' perception of co-articulated tones in tone sequences of tri-syllables and found that listeners compensate for the tone variations in their surface forms, due to the variability of their acoustic realizations depending on the tones of the preceding and following syllables. Tone identification is better in compatible contexts than conflicting contexts for Mandarin speakers. It is not known whether native Mandarin speakers perceive Mandarin lexical tones presented in compatible contexts and conflicting contexts created by two tone sequence in disyllables the same as the tri-syllables. As most of Mandarin Chinese words are disyllabic in modern times (Duanmu, 1999), it is necessary to study the perception of two tone sequence and see how the tonal contexts would influence their tone perception, which would be more constructive to apply in L2 acquisition .

Furthermore, previous studies (Stagray & Downs, 1993; Hallé et al., 2004; Xu et al., 2006; Xi et al., 2010) have shown that native Mandarin-speaking listeners perceive Mandarin lexical tones categorically. Stagray & Downs (1993) found that Mandarin speakers exhibit the decreased within-category sensitivity and demonstrate categorical perception of Mandarin tones compared with English speakers by judging variable tones at increments within the frequency range of a level tone-phoneme category. Taiwanese Mandarin-speaking listeners' perception of Mandarin lexical tones shows a higher degree of categorical perception ("quasi-categorical") than that of listeners of a non-tone language (French) with no exposure to tone language (Hallé et al., 2004).

Xu et al. (2006) explored the language experience (Mandarin speakers vs. American English speakers with no exposition to tone languages) and stimuli complexity (speech vs. non speech stimuli: harmonic tones) on the categorical perception of pitch direction. Results show evidence of strong categorical perception of speech stimuli for Mandarin but not English listeners. Xi et al. (2010) investigated the neurophysiological correlates of categorical perception of Mandarin speakers of Chinese lexical tones in Mandarin Chinese by using EEG (electroencephalogram). It provides strong neurophysiological evidence in support of categorical perception of lexical tones in Chinese. Relative to within-category deviants, the across-category deviants elicited larger MMN in the left recording sites, reflecting the long-term phonemic traces of lexical tones.

Concerning native Mandarin-speaking children's acquisition of perceiving Mandarin lexical tones. Liu et al. (2014) used mismatch responses (MMRs) to explore the dynamic changes of native Mandarin-speakers' speech perception abilities from early to middle childhood with an adult control group by using Tone 2 (LH) and Tone 3 (L). The study found that only the adult group demonstrated typical early mismatch negativity (MMN) responses, suggesting that the ability to discriminate specific speech cues in Mandarin lexical tone is a continuing process in preschool- and school-aged children.

To summarize, F0 height and F0 contour are the primary perceptual cues of native Mandarin-speaking listeners to perceive Mandarin lexical tones, while duration, amplitude and creaky voice are the secondary perceptual cues. Native Mandarin-speaking listeners have predominant confusion between Tone 2 (LH) and Tone 3 (L), as well as Tone 1 (H) and Tone 4 (HL). Preceding and following tones and syllable position influence the tone perception of native Mandarin-speaking listeners. Native Mandarin-speaking listeners perceive Mandarin lexical tones categorically and the ability to discriminate specific speech cues in Mandarin lexical tone is a continuing process among the native Mandarin-speaking children.

2.2 The Acquisition of Mandarin Lexical Tones by L1 English Speakers

Chinese is a typical tone language, where tone must be lexically specified, and English is a typical stress language, where rhythmic structure organizes an utterance (Yip, 2003). Tone perception is a function of the native language environment, and it is linguistically based (Mattock, 2006). Therefore, due to cross-language difference, L1 English learners of Chinese have particular difficulty in learning Mandarin lexical tones, the unique feature of Chinese language. Besides cross-language difference, Trofimovich et al. (2003) stated that the phonetic context also influences the acquisition of second-language sound segments of both children and adult learners. The present study focuses on exploring how the phonetic context influences the the non-native perception of Mandarin lexical tones.

In the following section both Mandarin lexical tone perception and production in isolation as well as in context by L1 English speakers are discussed.

2.1.1 The Perception of Mandarin Lexical Tones by L1 English Speakers

Many previous studies explored the cross linguistic difference in perceiving Mandarin lexical tones by L1 English speakers, including confusion pattern, perceptual cues, context effects, and the reasons that cause the disprency of the perception of Mandarin lexical tones between native Mandarin speakers and L1 English speakers .

Studies are interested in exploring the difference between the confusion pairs of L2 learners and native Mandarin speaking listeners both in isolation and in context. When perceiving Mandarin lexical tones in isolation, Huang & Johnson (2010) found that for Chinese listeners, the pairs of Tone 1 (H) and Tone 2 (LH), as well as Tone 2 (LH) and Tone 3 (L) were rated as significantly similar, indicating a role of phonology in determining perceptual salience. For American English listeners, the most distinctive pairs Tone 1 (H) and Tone 3 (L), as well as Tone 3 (L) and Tone 4 (HL) were rated as significantly different. The largest rating disparity lies with pairs Tone 1 (H) and Tone 4 (HL), as well as Tone 2 (LH) and Tone 4 (HL), which were more distinctive for the Chinese listeners. When perceiving tone in context, McGinnis (1996) tested beginning Chinese learners' (American English speakers) perception of Mandarin tone in single syllable, two-syllable sequence and three-syllable sequence and found that there is high rate confusion between Tone 2 (LH) and Tone 4 (HL), as well as Tone 1 (H) and Tone 2 (LH) within the first month of studying Mandarin Chinese. After the first month, the most confused pairs included Tone 2 (LH) and Tone 3 (L), as well as Tone 1 (H) and Tone 4 (HL). Lee et al. (2010) found that Tone 2 (LH) is consistently the most challenging tone to identify by L2 learners, when only partial acoustic information is available. In addition to the disparity of the confusion pairs between native Mandarin-speaking listeners and nonnative speakers, common confusion pairs are also found in previous studies, including Tone 2 (LH) and Tone 3 (L) as well as Tone 1 (H) and Tone 4 (HL), which may be due to the similarities in their tone F0 height, such as F0 onset and offset points, and contour shape (Gottfried & Suiter, 1997; Kiriloff, 1969; Miracle, 1989; Shen, 1989; Hao, 2012).

When presented in isolation, it is found that L2 learners do not perceive the perceptual cues of Mandarin tones, including its F0 height and F0 contour, the same way as native Mandarin-speaking listeners. First, F0 height is correlated with stress perception for L1 English learners of Chinese (White, 1981). White (1981) noted that since F0 is involved in signaling stress contrasts in English, Mandarin tones tend to be perceived as having various degrees of stress. Specifically, Tone 1 (H) and Tone 4 (HL) with high F0 are preferentially perceived as stressed, and Tone 3 (L) as unstressed. Second, L2 learners pay more attention to F0 height but not F0 contour in perceving Mandarin lexical tones (Gandour & Harshman, 1978; Gandour, 1983, 1984; Yang, 2010; Huang & Johnson, 2010). Gandour & Harshman (1978) used multidimensional scaling (MDS) and found that American English speakers place more weight on average pitch and extreme endpoints, which are the F0 height of onsets and offsets. Gandour (1983, 1984) found that English listeners give greater weight to the F0 height, but not the F0 contour. Yang (2010) and Huang & Johnson (2010) also confirmed that register plays an important role in the perception of tones by non-native speakers. Third, L2 learners pay more attention to the offset and onset F0 height (Huang, 2004; Gilber & Liu, 2013; Liu, 2013). Huang (2004) found that English listeners attend primarily to F0 onset and offset whereas Chinese listeners pay attention to contour as a whole by using isolated monosyllabic tones. Gilber & Liu (2013) and Liu (2013) narrowed down the extreme endpoints to the offset F0 height. The study explored English language listeners' ability to discriminate tone F0 contours to explore whether their just noticeable difference of F0 contour changes due to F0 shift position, F0 shift direction, F0 contour direction, or stimulus type by using a three-interval, forced-choice procedure. It is found that English listeners have significantly lower psychophysical thresholds for F0 shifts at the offset than at the onset, therefore, F0 shift position is the primary determinant. Fourth, L2 learners perceive contour tones different from level tones (Liu, 2013). Liu (2013) used level (Tone 1), rising (Tone 2), and falling (Tone 4) tones within or across Mandarin tone boundaries to measure thresholds of tone pitch discrimination of both young English and Mandarin Chinese native listeners. It is found that performance is equivalent between English and native Mandarin listeners for level tones, but significantly differ for rising (Tone 2) and falling (Tone 4) tones. In addition, Chang (2011) found that different from Mandarin speakers, L2 listeners give more perceptual weight to duration by identifying tone in isolation to differentiate Mandarin Tone 2 (LH) and Tone 3 (L), while the durational difference between Tone 2 and Tone 3 is the secondary difference for native Mandarin-speaking listeners (Massaro et al., 1985; Tseng, 1990; Whalen & Xu, 1992; Zeng, 2000; Kuo et al., 2008).

When presented in context, previous studies have tried to explore how tonal context and syllable position influence L2 learners' perception of Mandarin lexical tones in context. Broselow et al. (1987) tested American English listeners' perception of Mandarin tones in isolation, tone sequences of two and three syllables. The study found that Tone 4 is the most easily identified tone when presented in isolation, and in the final position of doublets and triplets and its identification declined dramatically in non-final positions. However, Fox and Qi (1990) found that there is no significant difference between perception in isolation and in the context condition for either language group. The study investigated whether context F0 influences tone perception for both native speakers of Mandarin and nonnative speakers in both isolation condition and in the context condition: tone sequences of two syllables. In the context condition, the initial syllable was either a Tone 1 (H) or Tone 2 (LH), while the onset F0 of the final syllable varied along a continuum from Tone 1(H) to Tone 2 (LH). Subjects were asked to rate the the tone of the final syllable in the pair, according to how closely it resembled Tone 1 (H) and Tone 2 (LH). No significant difference has been found, which may be due to the fact that the target syllable is not influenced by the tonal context of the preceding tone. Both the offset Tone 1 (H) and Tone 2 (LH) is high, therefore there is similar carry-over effect from the offset of the initial syllable to the onset of the final syllable. Chen (1997) also tested tone perception of monosyllables and longer utterances by American English learners of Chinese, however the tone is not evenly distributed in the test and how tonal context (the preceding and following tones) influence participants' perception is not incorperated. Gottfried & Suiter (1997) investigated Mandarin tone identification by native Mandarin listeners and non-native listeners when only partical acoustic information was available, including intact, silence-center, center-only and onset-only with or without the following syllable. The study found carry-over tonal co-articulation in the following syllable seems to help Mandarin listeners identify tones; English-speaking listeners are hindered by the presence of a following syllable. By using a similar method, it is further suggested by Lee et al. (2010) that the non-native listeners did not show evidence of using coarticulatory information. Bent (2005) tested native Mandarin and English participants (naïve listeners) discrimination of Mandarin tones in monosyllables and tri-syllable sequences. It was found that the English listeners' sensitivity to the Mandarin tone contrasts varied depending on the acoustic similarity of the pitch contours for contrasting tones and on whether the tones were presented in monosyllabic or tri-syllabic utterances. Mandarin listeners mostly attended to lexical tone targets, English listeners attended mostly to global aspects of the stimuli, the overall pitch contour and timing of pitch maximum, while the Mandarin listeners were highly sensitive to all contrasts.

As it has been presented, the tonal sequence of two has not been thoroughly examined in previous studies, though most of Mandarin Chinese words are disyllabic in modern times (Duanmu, 1999). He (2010) has tried to bridge this gap by comprehensively examining the American English speakers' ability to perceive Mandarin Chinese co-articulated tones in disyllabic words. The study tested the effects of tonal condition, tonal sequence, tonal context and syllable position. Tones in isolation were identified with higher accuracy rate than tones in sequence. Only Tone 1 (H) is found to be influenced by the compatible and conflicting tonal context. The perception Tone 2 (LH), Tone 3 (L) and Tone 4 (HL) was found to be affected by the syllable position, and the perception of final syllable position is better than initial syllable position. Though the study is comprehensive, the following aspects need to be incorporated to deepen our knowledge in L2 learners' perception of Mandarin tones in context. It still remains to be answered which tone combination in conflicting or compatible context is more problematic for beginning and more experienced L2 learners whose L2 aqcuisition time was more than one year, to acquire in the daily classroom learning environment. In addition, acoustic analysis of the stimuli need to be added; otherwise, we do not know how the F0 height and F0 contour is distorted in the tonal context and thereby contributes to the perception result. Also, it is unknown whether the acoustic variation in the stimuli and the correlation matrix between the F0 value of tone in monosyllables and disyllables would predict the L2 perception accuracy and error patterns. In addition, a native Mandarin-speaking control group should be added to the study; otherwise it is hard to know whether a native Mandarin-speaking listener would have different responses to matched stimuli. In the present study, a similar method has been adopted with improvement by adding acoustic analysis of the stimuli in the production study and adding a native Mandarin-speaking listener group as a control group in the perception study. The diversity of the stimuli has been improved by having two female speakers and perceiving different syllables in the sequence of two. It has been demonstrated that the Mandarin tone perception task with two female voices is more challenging than the task with a single female voice (Ning, 2014). Besides the accuracy rate and error pattern analysis, the sensitivity index (d') needs to be analyzed, which is calculated according to signal detection theory (Macmillan & Creelman, 2005). The sensitivity index (d') takes both the hit rate (the proportion of correct responses) and the false-alarm rate (the proportion of incorrect responses) into condiseration. It has been used in previous studies of Mandarin lexical tone perception (Chen & Massaro, 2008; Krenmayr et al., 2011). In an identification experiment, response bias is the tendency towards reporting one of the possible responses more often than others (Chen & Massaro, 2008; Krenmayr et al., 2011). Therefore, the hit rate (proportion of correct responses) is a biased index. Using d prime values to assess the discrimination performance for individual tones is suggested, as d prime values outperform the simple proportion of correct responses and can yield an unbiased esitimator of the underlying sensitivity. In addition, reaction time is included in the perception study.

Concerning the different behavioral results and perceptual cues between nonnative listeners and native Mandarin-speaking listeners in perceiving Mandarin lexical tones, models and studies have been trying to explain the discrepensy.

First, several models have proposed to explain the difficulty in non-native sound perception and production from the aspects of cross-language difference, which refers to how perceptually similar sounds are perceived in the native and second languages, including the second language theories of Perceptual Assimilation Model (PAM) (Best, 1995), Speech learning Model (SLM) (Flege, 1995) and Native Language Magnet Model (NLM) (Kuhl & Iverson, 1995) have explored this factor and indicate that L1 has a large impact on the acquisition of L2 production and perception. PAM (Best, 1995) focuses on non-native perception and posits that when perceiving non-native sounds, the non-native sounds will be mapped onto one or more native categories. if a non-native sound is assimilated to a native category, then primarily linguistic knowledge of the native language will be recruited. For non-native sounds that are assimilated as uncategoriezable speech sounds and nonassimilable non-speech sounds, listeners will process the sounds with less or no linguistic knowledge of the native language. SLM (Flege, 1995) accounts for how individuals learn – or fail to learn – producing and perceiving phonetic segments (vowels, consonants) in a second language. It emphasizes acquiring a new phonetic category, vowels and consonants, of nonnative sounds in terms of pronunciation. It is comparatively easy for learners to acquire new nonnative sounds and achieve native-like production, under the condition that the nonnative sound can be clearly distinguished from any L1 sound, otherwise learners' production of the nonnative sound will interfere with the pre-existing L1 phonological categories. NLM (Kuhl & Iverson, 1995) suggests that exposure to language early in life produces a change in perceived distances in the acoustic space underlying phonetic distinctions, and this subsequently alters both the perception and production. Specifically, phonetic prototype perceptually attracts surrounding stimuli. The region surrouding a good distance to the prototytpe of the category exhibits reduced sensitivity and percetual clustering. Perceptual distance between the prototype and its surrounding stumili is shrunk, while the region near the phonetic boundary is perceptually streched, which is attributable to langauge experience.

Second, studies have found to be aligned with the aforementioned models. L2 learners would be able to use their native prosodic categories to perceive foreign Mandarin tones (So & Best, 2011, 2014). The study found that both English and French speakers are able to perceptually categorize foreign tones into their intonational categories (i-Categories), and that categorizations are based on the contextual phonetic similarities of the F0 contours they perceived between Mandarin tones and their native i-Categories. The results showed that for English speakers, Tone 1 (H), Tone 3 (L), and Tone 4 (HL) are more often perceived as a statement; Tone 2 (LH) is more often perceived as a question. This is aligned with PAM model (Best, 1995), as non-native tonal category is mapped onto their intonational categories. According to the result from So& Best (2011, 2014), we would expect that Tone 2 is more perceptually marked, as it is different from other tones and perceived as a question. However, from the aforementioned perception result, we can see that this is not the case. Tone 2 (LH) is possibly the most challenging tonal category for L1 English speakers to perceive (Lee et al., 2010), as it is consistently confused with Tone 3 (L) (Gottfried & Suiter, 1997; Kiriloff, 1969; Miracle, 1989; Shen, 1989; Hao, 2012). It is more possible to categorize it as uncategoriezable speech sounds, therefore, the main discrepancy should not be caused by English intonational categories, but the similarities and the complex phonological relationships among uncategoriezable speech sounds. Another line of result is aligned with the NLM model (Kuhl & Iverson, 1995), that native Mandarin speakers have low sensitivty to the within category varianction, as these variations are attracted to the prototype, F0 value and F0 contour of Mandarin lexical tones in monosyllable perceptually. On the contracy, nonnative listeners are not exposed to lexical tones in their early life and their perceived distance of within category variations is not distorted in the acoustic space as the underlying tonal categories, therefore, nonnative listeners are sensitive to within categorical vairants. It is suggested that nonnative listeners, who have not acquired the phonetic boundaries between tones, may misperceive any within categorical phonetic variation as linguistically relevant (Stagray & Downs, 1993; Mattock, 2006; Huang & Johnson, 2010; Chang, 2011). Stagray & Downs (1993) found that compared with Mandarin speakers, English speakers maintained the within-category sensitivity and did demonstrate categorical perception of Mandarin tones by judging variable tones at increments within the frequency range of a level tone-phoneme category. Huang & Johnson (2010) explored the perceptual similarities of Mandarin lexical tones by native speakers of American English (AE) as well as Native Chinese speakers by using tones in isolation. It is suggested that without lexical tone categories in their lexicon, AE listeners may actually enjoy a perceptual advantage; that is, they may be able to detect subtle pitch differences, which may be missed by Chinese listeners' categorical perception of tone.

Other than the factors related to language specific perception, there are an additional two factors that influence L2 learners' perception of Mandarin lexical tones, including learning experience and music experience. First, learning experience contributes to the perception of Mandarin tones by L2 learners (Leather, 1990; Wang et al, 1999, 2003, 2006; He, 2010). Wang et al. (2006) summarized that many studies have shown that after short perceptual tone training, nonnative speakers of Mandarin improved both their perception and production of Mandarin tones (Leather, 1990; Wang et al, 1999, 2003). Specially, Wang et al. (1999) presented the result of training American listeners' identification of four Mandarin tones in isolation by using tone contrasts of natural stimuli in various phonetic contexts and spoken by various talkers. Their perception of Mandarin tones was improved after tone training in laboratory settings by using natural words produced by native Mandarin speakers. The trainees' identification accuracy revealed an average 21% increase from the pretest to the post-test, and the improvement gained in training was generalized to new stimuli with 18% increase and to new talkers and stimuli with 25% increase (Wang et al, 1999). He (2010) showed that experienced learners with more language learning experience outperformed inexperienced ones in perceiving Mandarin lexical tones in disyllabic words. In addition, music experience has been explored to see how it is relevant to the perception of nonnative suprasegmental patterns (Perrachione, 2007; Ning et al., 2014). Wong & Perrachione (2007) investigated the learning of nonnative suprasegmental patterns of Mandarin tones for word identification of English-speaking adults. Learning success was found to be associated with the learners' ability to perceive pitch patterns in a non-lexical context and their previous musical experience (Ning et al., 2014). Ning et al. (2014) suggested that discrimination of musical tones is correlated significantly with discrimination of Mandarin tones.

Besides behavioral studies, neurophysiological studies have been used to study the neural mechanisms of processing non-native Mandarin tones. Generally, different from native listeners, nonnative listeners processed Mandarin lexical tones in either right hemisphere or both hemispheres and the left lateralization is associated with improvements in perceiving non-native lexical tones (Gandour et al., 2000, 2004; Wong et al., 2004; Wang et al., 2001, 2003, 2004) . By using Positron Emission Tomography (PET) to study tone processing, it is suggested that phonological processing of native suprasegmental, specifically tones, occurs in the left hemisphere, while non-native listeners tend to use the right hemisphere (Gandour et al., 2000; Wong et al., 2004). In order to further explore the lateralization of lexical tone, Wang et al. (2001) examined dichotic perception of Mandarin tones by native and nonnative listeners. It is found that Mandarin tones are predominantly processed in the left hemisphere by native Mandarin speakers, whereas they are bilaterally processed by American English speakers with no prior tone experience. That is to say both hemispheres are engaged when listening to lexical tones for English-speaking listeners. By using the same technique, Wang et al. (2004) further showed that left-hemisphere dominance of processing Mandarin tone by native and proficient bilingual listeners, whereas nonnative listeners (Norwegian or American listeners) showed no evidence of lateralization, regardless of their familiarity with lexical tone. By using fMRI to examine brain activity of processing Mandarin lexical tones, Wang et al. (2003) found that the improvements in performance were associated with an increase in the spatial extent of activation in left hemisphere, after English speakers completed lexical tone training of learning Mandarin as a second language. Gandour et al. (2004) confirmed less activity for L1 English speakers in the left hemisphere than the Chinese group. It is proposed that the right hemisphere is sensitive to low-level acoustic processing and the left hemisphere is sensitive to high-level linguistic processing (as cited in Jongman et al., 2011).

To summarize, other than the commonly found conflusion pairs of Tone 2 (LH) and Tone 3 (L) as well as Tone 1 (H) and Tone 4 (HL), additional confusion pairs include Tone 2 (LH) and Tone 4 (HL), Tone 1 (H) and Tone 2 (LH) also found to be challenging to L2 learners. Concerning perceptual cues, F0 height, especially the F0 height of syllable offset, is the primary perceptual cue of L1 English learners of Mandarin to perceive Mandarin lexical tones. Both tonal context and syllable position seem to influence L2 learners' perception differently compared with native Mandarin-speaking listeners. The perception of contour tones, concerning Tone 2 (LH), Tone 3 (dipping), and Tone 4 (HL) is related more to the syllable position and the level tone (Tone 1) is more sensitive to the tonal environment for L2 learners. The disparity between native Mandarin listeners and L2 listeners may be due to language-specific perception, inlcuding the following reasons: L2 learners would be able to use their native prosodic categories to perceive foreign Mandarin tones; L2 learners would be more sensitive to the within categorical phonetic variations and misperceive them as different prosodic categories; L2 learners perceive Mandarin lexical tones mainly in right hemisphere, which is low-level acoustic processing. Learning experience and music experience positively correlated with the perception of Mandarin lexical tones.

2.1.2 The Production of Mandarin Lexical Tones by L1 English Speakers

Many previous studies have explored the cross linguistic difference and context effects in producing Mandarin lexical tones by L1 English speakers. The L2 acquisition of mandarin lexical tones and the correlation between perception and production has also drawn researchers' attention.

Many studies have explored the cross linguistic difference between L1 English and Mandarin Chinese speakers (White, 1981; Chen, 1977; Shen, 1989; Yang, 2010; Ning et al., 2014). By comparing the system of Chinese tones to that of English intonations, White (1981) found the major difference between Chinese tones and English intonations is that Chinese has significantly wider pitch range than English. Chen (1977) found that the average pitch range of the native Mandarin speakers is at least 1.5 times wider than that of the English speakers when they spoke their native languages. In addition, studies have tried to explore how the first language of L2 learners would influence their Mandarin lexical tone production. Shen (1989) found that Tone 1 (H) and Tone 4 (HL) are the most difficult for English speakers to produce. This is attributed to the fact that Tone 1 (H) and Tone 4 (HL) are most similar to the pitch patterns in English and thus more susceptible to L1 interference. In addition, comparing with native Mandarin speakers, the non-native speaker's production of Mandarin tones appears to be determined by the contour of the tones, but not determined by the register (Yang, 2010). Hao (2012) explored second language acquisition of Mandarin Chinese tones by English and Cantonese speakers. It is found that learners were significantly better at mimicking tones than at identifying or reading them, suggesting that the major difficulty that learners faced in acquiring Mandarin tones was associating pitch contours with discrete tonal labels. Hao (2012) suggested that factors other than learners' L1 background may also play a significant role in explaining difficulties in acquisition of certain L2 sounds, such as acoustic similarity and complex phonological relationship. Ning et al. (2014) suggested that the F0 contours of nonnative speakers were more affected quantitatively by the amplitude and direction of pitch perturbations, suggesting less stable internal tone models than Mandarin speakers.

Concerning producing Mandarin lexical tones in context, many studies are interested in exploring the errors that are made by L1 English speakers. White (1981) found that American English learners' tonal errors mainly occur in sentences or in combinations of words, not in isolated syllables. Through word-level acoustic analysis, Miracle (1989) showed that American learners made roughly the same amount of errors across the four tones by analyzing the words produced in topic position. It is suggested that although there is no significant difference, Tone 2 (LH) might be particularly difficult for American learners. Chen (1997) analyzed the tonal errors in sequence of American English learners of Chinese, including monosyllables, disyllables, 10 syllables, and longer sequences, and found that there were two production problems, including alien level tones and non-fully realized contour tones. Guo & Tao (2008) reported the finding on the developmental stages of tone production at sentence level by American students in their first year of learning Mandarin Chinese and confirmed the difficulty of Tone 3 (L) in near-natural conversations. He & Wayland (2010) tested the effects of language learning experience, tonal environment, tonal context and syllable position on American English speakers' ability to produce Mandarin Chinese coarticulated tones in disyllabic words. It was found that with increased experience, production of coarticulated tone becomes more resistant to such phonological and phonetic factors as tonal contexts and syllable position. In addition, American speakers' production of Mandarin coarticulated in both compatible tonal environment and conflict tonal environment was equally accurate and American speakers produce tones more accurately in final syllable position than in initial syllable position. It was found that tonal environment only affected Tone 1 (H) production and syllable position affected Tone 2 (LH) and

Tone 4 (HL) production accuracy. Neither tonal environment nor syllable position has an influence on the accuracy rate of production of Tone 3 (L).

Pedagogical studies have been done to explore teaching methods to improve L2 learners' acquisition of Mandarin lexical tones, including acquisition order (Yue-Hashimoto, 1986; Chen, 1997), and training methods (Yue-Hashimoto, 1986; Lu, 1992; Shen, 1989; Chang, 2008; Wang et al., 2003; Wang et al., 2006). Concerning acquisition order, Yue-Hashimoto (1986) found that the tonal acquisition order for native Mandarin speakers and L2 learners of Mandarin Chinese is similar: Tone 1 (H) is the easiest to acquire, followed by Tone 4 (HL) , Tone 2 (LH) and Tone 3 (L). This is further confirmed by Chen (1997), which analyzed the tonal errors in sequence of American English learners of Chinese, including monosyllables, disyllables, 10 syllables, and longer sequences. Concerning training methods, Yue-Hashimoto (1986) suggested that Mandarin tones are best instroduced in a bi-syllable environment. Since the bi-syllabic compounds are the most prevalent in Chinese, and the tone values change slightly in a compound environment, compared to in an isolated syllable environment. Lu (1992) also suggests that teachers should introduce bi-syllables first and tonal perception is more crucial in the beginning stages than tonal production. Shen (1989) recommends that three pitch levels (low, mid and high) should be used instead of the traditional five and that the emphasis should be on the F0 height instead of the F0 contour. Chang (2008) proposed a design for a beginner's course in Chinese pronunciation focusing on mastering Mandarin tones. It is suggested that half third tone should be introduced before full third tone; Mandarin instructors should pay more attention to perception ability in the beginning stage; most of the tonal practice should start with bi-syllable phrases so that students can get the sense of Tone Sandhi more easily in the later stage. In laboratory setting, Wang et al. (2003) revealed significant tone production improvement after perceptual training of Mandarin tone contrasts, which sugests that tone contrasts gained perceptually could transfer to production among L1 English learners of Mandarin Chinese. Tone pair confusion analysis showed that Tone 1 and Tone 4 were most resistant to improvement; while Tone 2 and Tone 3 were improved greatly after training. Therefore, it is suggested that pitch height and pitch contour are not mastered in parallel. Their production of pitch contour improved more than the pitch height. Wang et al. (2006) summarized that many studies have shown that after short perceptual tone training, nonnative speakers of Mandarin improved both their perception and production of Mandarin tones (Leather, 1990; Wang et al. 1999, 2003).

Concerning the correlation between Mandarin lexical tone production and perception in context, many studies suggested that English speakers' perception and production ability of Mandarin lexical tones is not significantly correlated (Chen, 1997; Bent, 2005; He, 2010). Chen (1997) suggested that tonal production and perception did not seem to correlate by testing the perception and production of monosyllables and longer utterances. Bent (2005) examined the perception and production of Mandarin lexical tones in monosyllables and tri-syllable sequences, by testing native Mandarin and English participants (naïve listeners)' discrimination and imitation of Mandarin tones. It is found that listeners' perception and production abilities were not significantly correlated. He (2010) tested the effects of language learning experience, tonal condition, tonal sequence, tonal context and syllable position on American English speakers' ability to perceive and produce Mandarin Chinese coarticulated tones in disyllabic words. It is suggested that relationship of perception and production of isolated tones for American learners was still not clear. However, in the laboratory training setting, a positive correlation between Mandarin lexical tone perception and production of American English speakers is observed (Wang et al., 1999, 2003). It seems that the mismatch between the nonnative perception and production exists among two Tone pairs, including Tone 2 (LH) and Tone 3 (L), as well as Tone 1 (H) and Tone 4 (HL). Tone 2 (LH) and Tone 3 (L) confusion is more of a difficulty in perception (Gottfried & Suiter, 1997; Kiriloff, 1969; Miracle, 1989; Shen, 1989; Hao, 2012), and Tone 1 (H) and Tone 4 (HL) is more of a persistent problem in production (Shen, 1989; Wang et al., 2003). According to SLM (Flege, 1995), the prerequisite of being able to produce and acquire a new phonetic category of sound is that the nonnative sound can be clearly distinguished from L1 sounds. In the case of perceiving and production nonnative Mandarin lexical tones, the link between perception and production can be partially explained by SLM. On one hand, Tone 1(H) and Tone 4 (HL) are most similar to the pitch patterns in English (Shen, 1990), and they both are perceived as English statement intonation (So & Best, 2011, 2014). It is possible that the L1 English learners' production difficulty of acquiring Tone 1 (H) and Tone 4 (L) is due to the influence from the pre-existing L1 English phonological categories, as possibly they have great difficulty in distinguishing Tone 1 (H) and Tone 4 (HL) from English statement intonation. On the other hand, Tone 2 (LH) is mainly perceived as a question and Tone 3 (L) is mainly perceived as a statement (So & Best, 2011, 2014), which has been found that L1 English learners of Chinese are having less difficulty in producing Tone 2 (LH) and Tone 3 (L) (Shen, 1990, Wang et al., 2013). Counter examples have been found that Tone 2 (Miracle, 1989) and Tone 3 (Guo & Tao, 2008) is harder to produce correctly in context, which may be due to the similarity between the tone categories and complex phonological relationships.

To summarize, both L2 learners' L1 backgound and the acoustic similarity and complex phonological relationship would contribute to explain the difficulties in acquiring Mandarin lexical tones. When producing tone in context, Tone 2 (LH) and Tone 3 (L) may still be harder to produce than other tones (Miracle, 1989; Guo & Tao, 2008), although in controlled laboratory settings, the tone contrasts gained perceptually significantly improved Tone 2 (LH) and Tone 3 (L) production for non-native speakers (Wang et al., 2003). Both the tonal context and syllable position influence the L2 production of Mandarin lexical tones (He & Wayland, 2010). The production errors include alien level tones and non-fully realized contour tones (Chen, 1997). Studies have been trying to explore the effective teaching methods to improve L2 production of Mandarin lexical tones, including acquisition order (Yue-Hashimoto, 1986; Chen, 1997), and training methods (Yue-Hashimoto, 1986; Lu, 1992; Shen, 1989; Chang, 2008). Although in the laboratory setting, Wang et al. (2003) suggests that tone contrasts gained perceptually could transfer to production among L1 English learners of Mandarin Chinese, many studies suggested that English speakers' perception and production ability of Mandarin lexical tones is not significantly correlated (Chen, 1997; Bent, 2005; He, 2010). It is possible that the mechnism causing the difficulty in acquring Tone 1 (H) and Tone 4 (HL), as well as Tone 2 (LH) and Tone (L) 3 is different. The difficulty in acuiqring Tone 1 (H) and Tone 4 (HL) is more related to cross linguistic difference, and Tone 2 (LH) and Tone 3 (L) is more related to the similarity between the tone categories and complex phonological relationships.

The following two studies in Chapter 3 study I: the production of mandarin lexical tones in context by l1 chinese speakers and Chapter 4 study II: the identification of mandarin lexical tones in context are designed to answer how tones produced in context would influence the L2 tone identification for both beginning and more experienced L2 learners. These two studies will help us to deepen our knowledge of tone acquisition in the daily classroom learning environment, as tone production and perception is intimately related to using Mandarin in everyday life, which is highly influenced by the tonal contexts.

CHAPTER 3 STUDY I

THE PRODUCTION OF MANDARIN TONES IN CONTEXT BY L1 CHINESE SPEAKERS

The following study adopted a similar experimental design as Xu (1997) in order to examine the carryover effects and the anticipatory effect to explore the native Mandarin speakers' production of tonal sequence of two in Mandarin Chinese. In addition, the concept of the compatible and conflicting context from Xu (1993, 1994) was incorporated as one of the factors. The present study offered a combined view to systematically examine native tonal co-articulation concerning tone sequence of disyllabic words in Mandarin.

3.1 Research questions

- How does the F0 height and F0 contour of tone sequence of two deviate from the tone in monosyllable of Mandarin lexical tones?
- How does the placement of a tone in either the initial or final syllable position affect the F0 height and F0 contour of tone sequence of two of Mandarin lexical tones, specifically concerning carry-over effects and anticipatory effects?
- How do the compatible and conflicting contexts influence the F0 height and F0 contour of Mandarin lexical tones?

3.2 Hypothesis and predictions

- The tone of the initial syllable and final syllable should be different from the tone in monosyllable in terms of both F0 height and F0 contour.
- It is expected that the F0 height and F0 contour deviateion of tone in the initial syllable position is mainly caused by a carry over effect: the onset of a tone is assimilated into the offset value of the previous tone; and the deviation of tone in

the final syllable position is mainly caused by anticipatory effect: the onset value of a tone dissimilate to the F0 value of a preceding tone.

- The magnitude of both carry over effect and anticipatory effect is smaller in compatible context than conflicting context.

3.3 Methods

3.3.1 Informants

Two female informants from Beijing produced both the monosyllabic and dissyllabic Mandarin tones. They both are from Beijing, and have never moved outside of Beijing before studying at IU. Their first language is Mandarin, which they speak at both home and school. They do not speak any dialect. Their second language is English. During the self-assessment, they rated themselves to be very comfortable communicating in Mandarin Chinese. All methods reported here were approved by the Institutional Review Board at Indiana University – Bloomington and all participants provided written informed consent.

3.3.2 Stimuli

Only the vowel /a/ was used in the stimuli, as it can readily be combined with different consonants to form words carrying all four Mandarin tones. No other vowels were used since F0 patterns vary among different vowels (Howie, 1976), which may affect non-native listeners' Mandarin tone perception. Two sonorants - the nasal /m/ and the glide /j/ - were used, as they are good tone carriers. In addition, the two syllables /ma/ and /ja/ are both possible words when carrying any of the four Mandarin lexical tones and with high familiarity to native Mandarin speakers.

The stimuli contained monosyllables, disyllables, and disyllables in carrier phrases. In total there were 40 stimuli, including 8 monosyllabic stimuli and 32 disyllabic stimuli. The monosyllabic stimuli included 8 tokens (4 tones×2 syllables), as shown below (See Table 3).

Table 3: Monosyllabic stimuli

Simplified Character	Syllable	Gloss
妈	Tone 1 (H) /ma/	mother
麻	Tone 2 (LH) /ma/	hemp
马	Tone 3 (L) /ma/	horse
骂	Tone 4 (HL) /ma/	scold
压	Tone 1 (H) /ja/	press
牙	Tone 2 (LH) /ja/	tooth
雅	Tone 3 (L) /ja/	elegant
讶	Tone 4 (HL) /ja/	surprised

Two sets of syllables, including /ma.ja/ and /ja.ma/, composed the disyllabic stimuli, which included two disyllabic combinations of 16 tokens (4 initial tones × 4 final tones). In total, the disyllabic stimuli included 32 tokens (4 initial tones × 4 final tones × 2 syllables) (See Table 4). The dissyllabic tokens were all nonsense combinations. Using nonsense sequence is meant to limit the lexical frequency influence on the following tone identification tasks, when these tokens serve as stimuli in Chapter 4. Tseng & Lee (2013) suggests that the exposure frequency seemed to play a role in the recognition of tones.

Table 4: Disyllabic stimuli

Tone 1 Tone 1 /ma.ja/	Tone 2 Tone 1 /ma.ja/	Tone 3 Tone1 /ma.ja/	Tone 4 Tone 1 /ma.ja/
Tone 1 Tone 2 /ma.ja/	Tone 2 Tone 2 /ma.ja/	Tone 3 Tone2 /ma.ja/	Tone 4 Tone 2 /ma.ja/
Tone 1 Tone 3 /ma.ja/	Tone 2 Tone 3 /ma.ja/	Tone 3 Tone3 /ma.ja/	Tone 4 Tone 3 /ma.ja/
Tone 1 Tone 4 /ma.ja/	Tone 2 Tone 4 /ma.ja/	Tone 3 Tone4 /ma.ja/	Tone 4 Tone 4 /ma.ja/
Tone 1 Tone 1 /ja.ma/	Tone 2 Tone 1 /ja.ma/	Tone 3 Tone1 /ja.ma/	Tone 4 Tone 1 /ja.ma/
Tone 1 Tone 2	Tone 2 Tone 2	Tone 3 Tone2	Tone 4 Tone 2

/ja.ma/	/ja.ma/	/ja.ma/	/ja.ma/
Tone 1 Tone 3 /ja.ma/	Tone 2 Tone 3 /ja.ma/	Tone 3 Tone3 /ja.ma/	Tone 4 Tone 3 /ja.ma/
Tone 1 Tone 4 /ja.ma/	Tone 2 Tone 4 /ja.ma/	Tone 3 Tone4 /ja.ma/	Tone 4 Tone 4 /ja.ma/

The 16 disyllabic tone combinations can be broadly divided into two categories: compatible and conflicting contexts, according to the F0 height of the onset and offset (Xu, 1993, 1994; He, 2010) (See Table 5).

Table 5: The tonal context of tone sequence of two syllables: the conflicting and compatible contexts

Tone combination	1 1	1 2	1 3	1 4
F0 contour				
F0 value	H H	H LH	H L	H HL
Tone combination	2 1	2 2	2 3	2 4
F0 contour				
F0 value	LH H	LH LH	LH L	LH HL
Tone combination	3 1	3 2	3 3	3 4
F0 contour				
F0 value	L H	L LH	LH L	L HL
Tone combination	4 1	4 2	4 3	4 4
F0 contour				
F0 value	HL H	HL LH	HL L	HL HL

Note: Straight lines (▬) indicate compatible tonal combinations and dotted lines (= = = =) indicates conflicting tonal combinations, and the vertical lines indicate the syllable boundary of disyllables. Because of Tone 3 sandhi rule: the first Tone 3 (L) changes to LH, which is similar with Tone 2 (LH), when preceding another Tone 3.

3.3.3 Procedure

Please see Appendix I for the monosyllabic reading list and the disyllabic reading list. One speaker first produced the monosyllables /ma/ and /ja/ with the four lexical tones in isolation three times. Then the speaker was asked to produce all 16 /ma.ja/ and /ja.ma/ sequences three times as naturally as possible, i.e. with no pause between the two syllables. The other speaker was asked to produce the same sets of stimuli in a counterbalanced order. Background information was gathered at the end. The tokens were recorded on a laptop using a Praat software package (Boersma, 2001), a free software package for the acoustic analysis of speech.

The disyllables, though nonsense words, were read as naturally as possible, i.e. with no pause between the two syllables. The stimuli were recorded at 44100Hz on a laptop using a Praat software package (Boersma, 2001), and a high-quality stereo headset from Microsoft Life Chat LX-3000.

Two native speakers of Mandarin Chinese from Beijing identified the tones. Only good exemplars of the target tones were used in the data analysis; bad exemplars were deleted and replaced with newly recorded ones, which were also examined by the two native speakers to ensure the validity of the tone sequences.

3.4 Data analysis

Only the second and third repetitions were used in the data analysis. There were a grand total of 160 tokens (40 stimuli ×2 speakers × 2 repetitions), used in the experiment. Specifically, Monosyllables: 32 tokens (4 tones × 2 syllables × 2 informants × 2 repetitions); disyllables:128 tokens (4 initial tones × 4 final tones × 2 syllables × 2 informants × 2 repetitions).

F0 contour for each tone was divided with 16 equal length F0 segments. The time scale was equalized by all the curves by using the following method: the F0 value of the 15 cutting points plus the starting and the ending points of F0 value were taken from the obtained F0 curves by using Praat (Boersma, 2001) and further analyzed and plotted by using R (R Development Core Team, 2008), a free software environment for statistical computing and graphics. Several F0 measurements obtained from F0 curves are further discussed in the statistical analysis, including

F0 values at onset, one quarter , one half , and three quarters, and offset, and duration. For tone sequences of two syllables, the initial syllable and final syllable were spliced clearly with sufficient information from nasality (/ja.ma/) and formant transition (/ma.ja/).

3.5 Results

3.5.1 Monosyllables

Though the main focus of this study is on the contextual influence of Mandarin tones, the production results from monosyllables are reported here to serve as the baseline.

After the visual representation of F0 for each syllable “/ma/” and “/ja/” was constructed, no great differences were found between the two speakers. Therefore, the average F0 value of the two speakers, two syllables and two repetitions was obtained as the abstraction of the base value of the four tones in the study (See Figure 3). The F0 patterns of tones produced in isolation directly reflect the canonical forms of the tones of monosyllables, as described by 5-level scale (Chao, 1930) and F0 height and F0 contour (Howie, 1976) (See Table 1), however several differences are worth mentioning. For Tone 3, the onset and offset are both very low, which is very different from the canonical description as mid-falling-rising tone. Tone 3 (L) is the lowest tone compared with other tones. This abstract representation supports previous studies that Tone 1 (H) and Tone 4 (HL) share similar high onset value, Tone 1 (H) and Tone 2 (LH) share similar high offset value, Tone 2 (LH) and Tone 3 (L) share similar low onset value, and Tone 3 (L) and Tone 4 (HL) share similar low offset value.

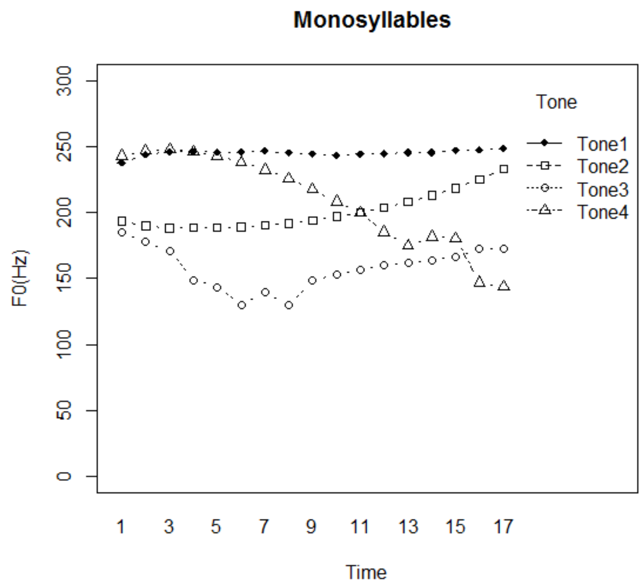


Figure 3: The abstraction of Mandarin tones in monosyllable

Therefore, the F0 patterns of tones produced in isolation, can be described the same as the F0 patterns when produced in context. As has been mentioned in literature review, the four underlying tones in connected speech of Mandarin Chinese are Tone 1 (H), Tone 2 (LH), Tone 3 (L), and Tone 4 (HL) (Shih, 1987; Xu, 1997). See the following table, for the F0 value at onset, one quarter, one half, three quarters, and at offset of the Mandarin lexical tones in monosyllables based on the stimuli in the present study.

Table 6: F0 value (Hz) of the tones in monosyllables

Target Tone	F0 onset	One quarter	One half	Three quarters	F0 offset
Tone 1 (H)	237.77	245.80	244.19	245.42	248.45
Tone 2 (LH)	193.40	188.62	194.19	207.94	224.99
Tone 3 (L)	184.88	143.55	148.52	161.62	172.86

Tone 4 (HL)	242.89	243.07	217.37	175.13	143.70

Spearman rank correlation (Baayen, 2008), a distribution-free, non-parametric correlation test, was performed to examine the degree of similarity between two rankings between the average F0 value of the tones in monosyllable, and further to assess the significance of the relation between them. The assessment can reveal the extent of statistical dependence between pairs of F0 values for the 15 cutting points, thereby assessing the similarity between six pairs of Mandarin Tones (Tone 1 (H) vs Tone 2 (LH), Tone 1 (H) vs Tone 3 (L), Tone 1 (H) vs Tone 4 (HL), Tone 2 (LH) vs Tone 3 (L), Tone 2 (LH) vs Tone 4 (HL), and Tone 3 (L) vs Tone 4 (HL)). From the following correlation matrix, we can see that the rankings of F0 value of tones in monosyllable is dictinct from other tones, except for that Tone 2 (LH) is significantly and negatively correlated with Tone 4 (HL). There is moderate correlation between the rankings of F0 value of Tone 2 (LH) and Tone 3 (L) (rho=0.4), however it is not statistically significant.

Table 7: Correlation matrix of F0 value of tone in monosyllables

Tone	Tone 1 (H)	Tone 2 (LH)	Tone 3 (L)	Tone 4 (HL)
Tone 1 (H)	1.00 ($p<0.001$)	0.18	0.00	-0.29
Tone 2 (LH)	0.18	1.00 ($p<0.001$)	0.40	-0.96 ($p<0.001$)
Tone 3 (L)	0.00	0.40	1.00 ($p<0.001$)	-0.19
Tone 4 (HL)	-0.29	-0.96 ($p<0.001$)	-0.19	1.00($p<0.001$)

Note: Probability values adjusted for multiple tests.

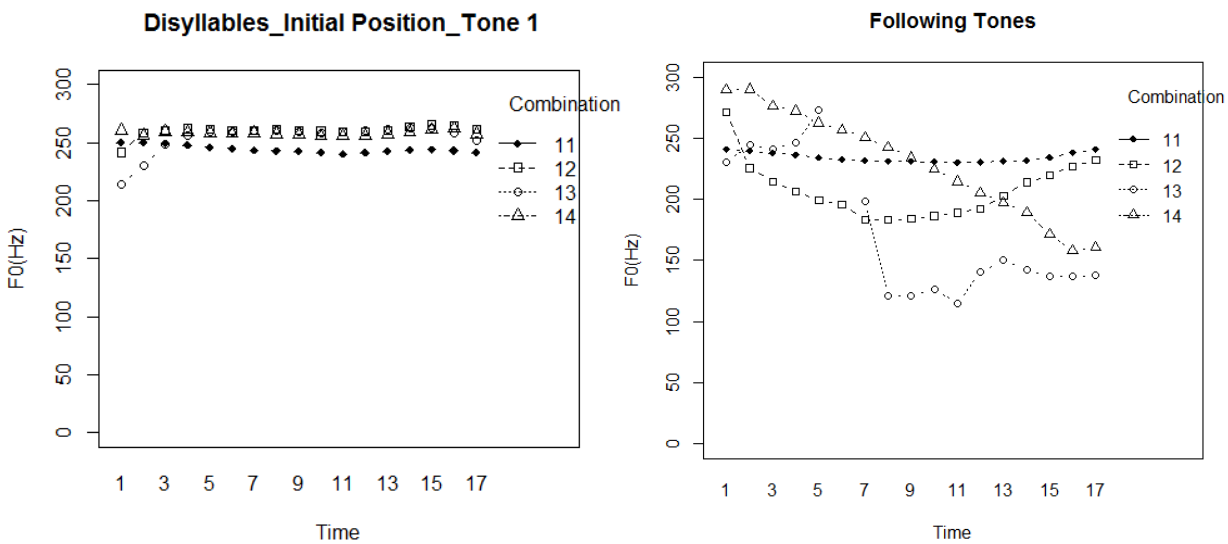
In terms of duration, on average Tone 4 (HL) is the shortest (417 ms), and Tone 3 (L) is the longest (568 ms). Tone 1(H) and Tone 2 (LH) have intermediate duration with Tone 2 (LH) (543 ms) longer than Tone 1 (H) (524 ms). These duration patterns match those found in previous studies (Lin, 1988; Xu, 1997).

3.5.2 Disyllables

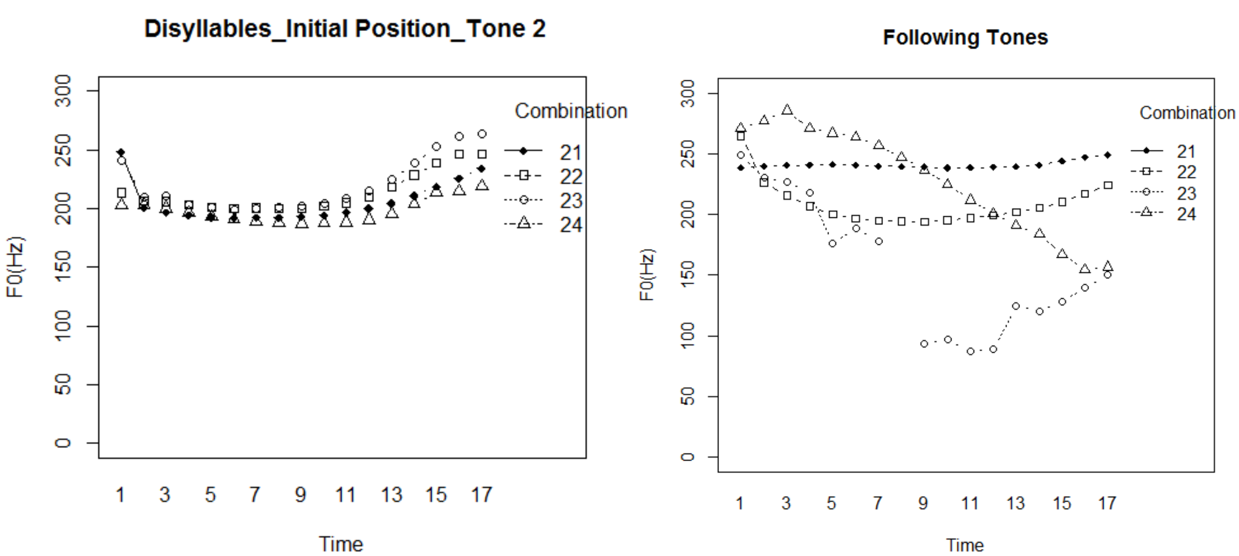
The factors that contribute to the contextual tonal variations are related to both position (carry over effect and anticipatory effect) and tonal combination (compatible and conflicting contexts). The factors are examined individually, as analyzed below.

3.5.2.1 Anticipatory effects: influences from the onset of final syllables to the offset of the initial syllable

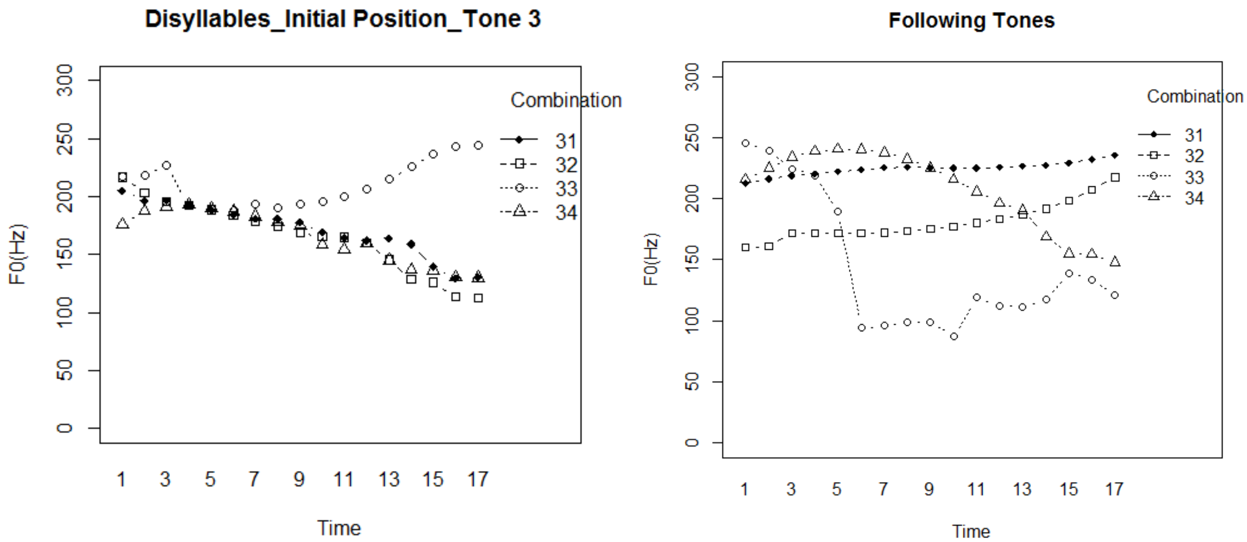
Figure 4 shows the F0 contour variation due to the influence of the onset of the tones in the final syllable /ma.ja/ and /ja.ma/ sequences. The same tones remain in the initial syllables, while followed by the four different tones. Except for the Tone 3 and Tone 3 sequence, all the other initial tones conserve the contour and pitch height of tones produced in isolation. The Tone 3 and Tone 3 sequence was excluded from the figure and was separately analyzed, due to the phonological tone sandhi that occurs when Tone 3 is followed by another Tone 3: the first Tone 3 (L) changes to LH, which is similar to Tone 2 (LH). From the figure, we can see that the anticipatory effect is little and the offset of the initial syllable is not significantly influenced by the onset of the final syllable. Similar phenomenon has been confirmed as described in Xu (1997), that when the initial syllable is followed by a tone with low onset (i.e. Tone 2 (LH) and Tone 3 (L)), its F0 contour is somewhat higher than when it is followed by a tone with a high onset (i.e. Tone 1 (H) and Tone 4 (HL)), therefore the property of the anticipatory effect is dissimilatory. This effect can be clearly seen when the tone of the initial syllable is Tone 1 (H), Tone 2 (LH) and Tone 4 (HL).



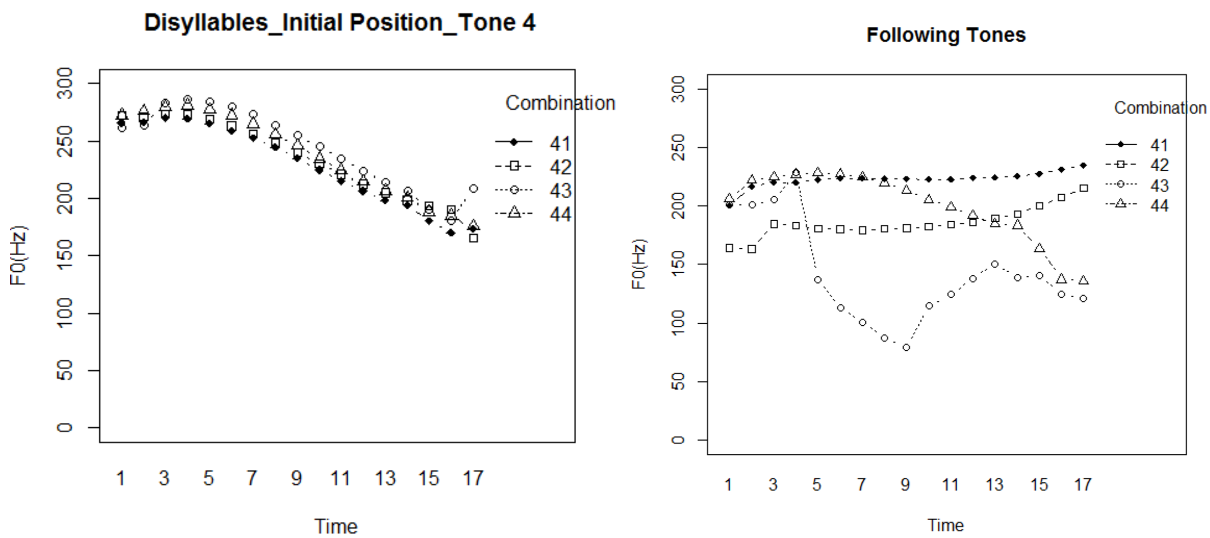
(a)



(b)



(c)



(d)

Figure 4: Influences from the onset of final syllables to the offset of the initial syllables

To examine the anticipatory effect, two sets of two way ANOVA (analysis of variance) tests and one set of three way ANOVA test were conducted. The first set was a two-way ANOVA with just the main effects. The independent variables were the four Mandarin tone in the initial syllable (Tone 1(H), Tone 2 (LH), Tone 3 (L), and Tone 4 (HL)) and the onset value of the final syllable in the /ma.ja/ and /ja.ma/ sequences (H: high and L: low). The dependent variables were the F0 values at onset, one quarter, one half, and three quarters, and at the offset of the initial syllable in the /ma.ja/ and /ja.ma/ sequences. The second set was a two-way ANOVA with the main effects and two-way interactions. However, none of the two-way interactions was significant, and they were not listed here. For the three-way ANOVA test, contexts (compatible and conflicting) were included as one of the factors and the effect was not significant. The following table shows the main effects of the two-way ANOVA test. The effect of the onset of the Final tone (non- target) was not significant. Due to the phonological tone sandhi that occurs when Tone 3 is followed by another Tone 3 (See Figure 4 (c)), the sequence Tone 3 and Tone 3 is not included in the analysis.

From Table 8, we can see that there is no anticipatory effect from the onset F0 value of the final syllable to the F0 value of the initial tone (target).

Table 8: ANOVA results for the effects of F0 onset value of the tone in the final syllable position to the F0 tone value at five positions in the initial syllable of the sequence /ma.ja / and /ja.ma/ sequences

position						
		0	0.25	0.5	0.75	1
Effect	df	F	F	F	F	F
Initial tone (target)	3,10	8.10 ($p<0.01$)	212.76 ($p<0.001$)	116.78 ($p<0.001$)	75.41 ($p<0.001$)	54.19 ($p<0.01$)

Onset of the Final tone (non-target)	1,10	0.08 (p=0.78)	3.71 (p=0.08)	3.42 (p=0.09)	3.75 (p=0.08)	1.61 (p=0.23)

In addition, Tone 3 in the initial sylable in the tone sequence Tone 3 and Tone 3 (L) was compared with Tone 2 (LH) in the initial syllable in the tone sequence Tone 2 (LH) and Tone 3 (L). It was found that F0 value of Tone 3 (210 Hz on average) in the initial syllable in the Tone 3 and Tone 3 sequence was lower than Tone 2 (219 Hz on average) in the initial syllable in the Tone 2 and Tone 3 sequence, although there was no significant difference between the F0 value of these two tones.

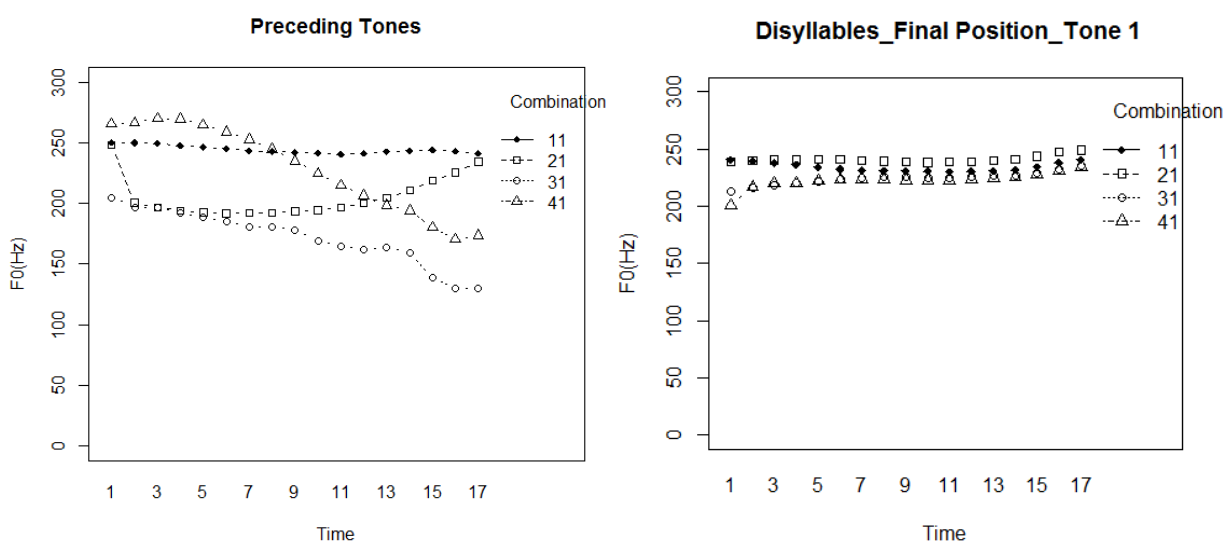
In terms of duration, a two-way ANOVA was performed on the duration of the initial syllable, with four Mandarin tones in the initial syllable and four Mandarin tones in the final syllable as independent variables. The durational variation due to neither the tone of the initial or the final syllable was significant. The mean durations of the initial syllables of the four tones were listed in the following table. Compared with the duration of the tones in monosyllable, the mean durations of the tones in the initial syllables was shorter in general (see Table 9). Tone 2 (LH) was the longest tone in initial position and Tone 4 (HL) was the shortest.

Table 9: Duration of the four Mandarin tones in the initial syllable

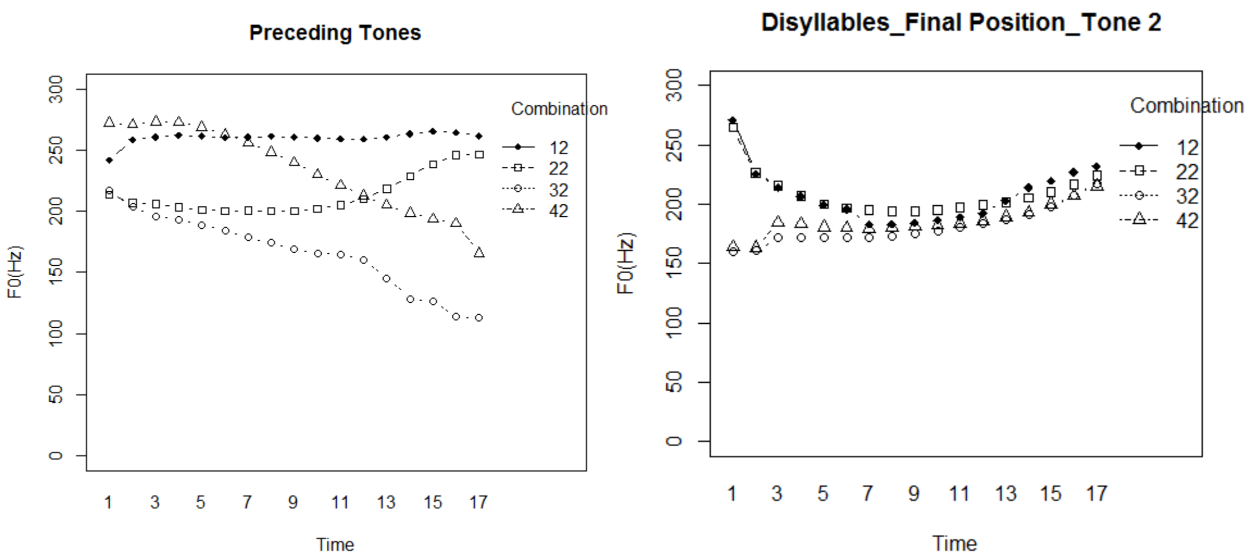
Tone 1 (H)	Tone 2 (LH)	Tone 3 (L)	Tone 4 (HL)
404 ms	428 ms	390 ms	385 ms

3.5.2.2 Carry over effect: Influences from the offset of first syllables to onset of the second syllables

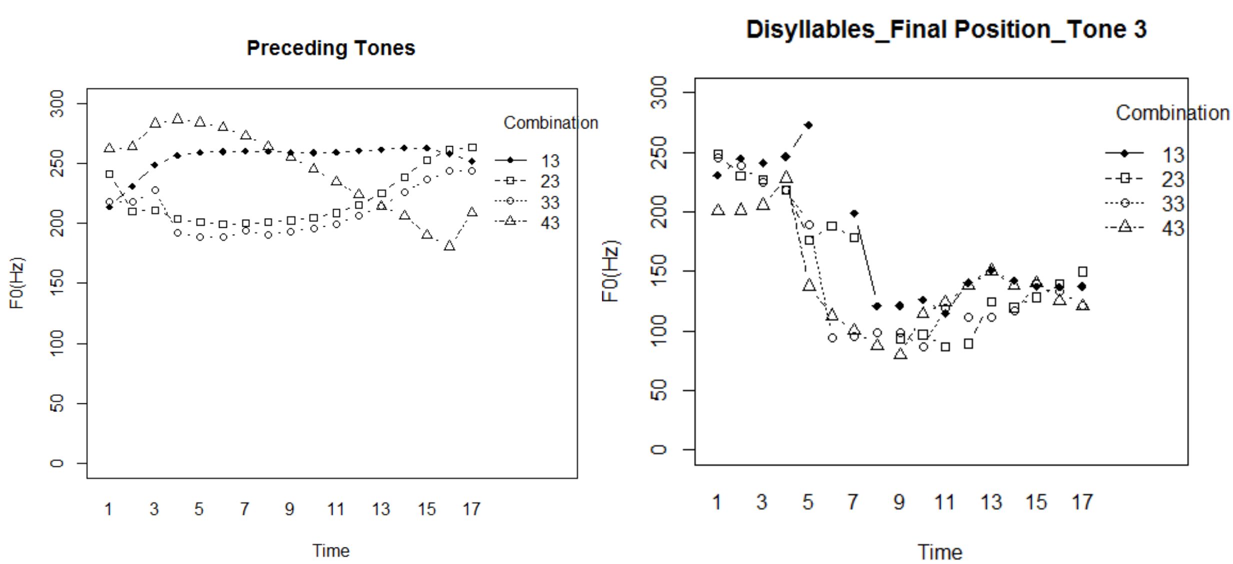
Figure 5 shows the F0 contour variation due to the influence of the offset of the initial syllable tones in the /ma.ja/ and /ja.ma/ sequences. The same tones remain in the final syllables, while preceded by the four different tones. Generally, the final tones were heightened by the initial Tone 1 (H) and Tone 2 (LH), as they both had a high offset, and lowered by the initial Tone 3 (L) and Tone 4 (HL), as they both had a low offset. Except for the Tone 3 and Tone 3 sequence, Tone 3 (L) was heightened by the initial Tone 3 due to the phonological variation (Chao, 1968). Also, from the figure we can see that Tone 3 in the final syllable position is different from Tone 3 (L) in the initial position in terms of F0 value and F0 contour; therefore it can be best described as dipping tone.



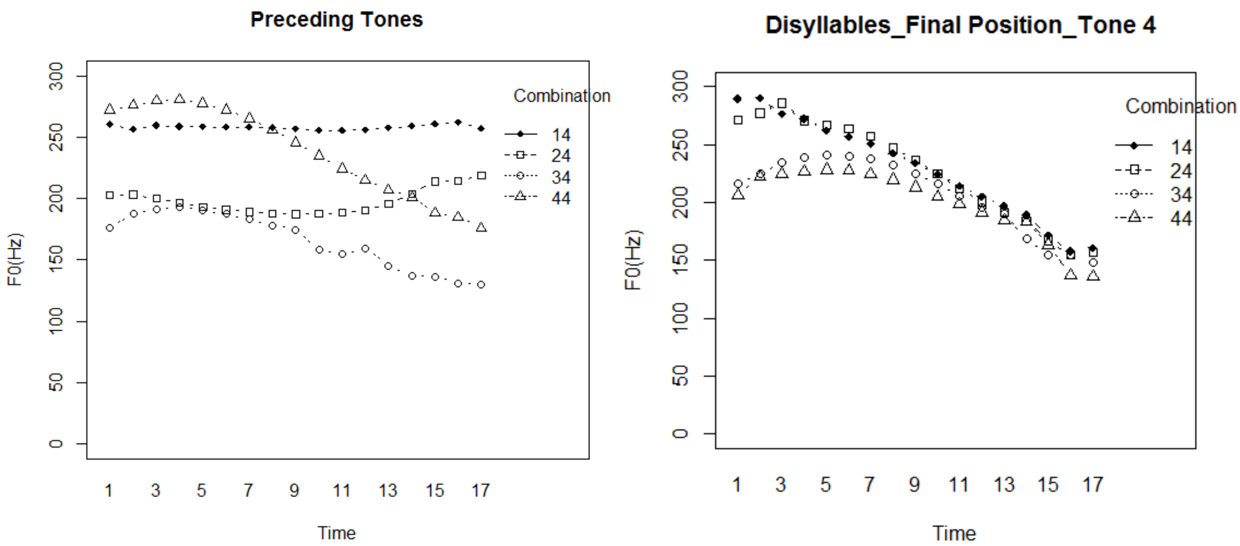
(a)


Preceding Tones
Disyllables_Final Position_Tone 2
F0(Hz)
Time
Combination
12
22
32
42
0
50
100
150
200
250
300
1
3
5
7
9
11
13
15
17

(b)


Preceding Tones
Disyllables_Final Position_Tone 3
F0(Hz)
Time
Combination
13
23
33
43
0
50
100
150
200
250
300
1
3
5
7
9
11
13
15
17

(c)



(d)

Figure 5: Influences from the offset of the initial syllables to the onset of the final syllables

To examine the anticipatory effect, one set of three-way ANOVA (analysis of variance) test with the main effect and two-way interactions were conducted. The independent variables were the four Mandarin tone in the final syllable, and the offset F0 value of the tones in the initial syllable in the /ma.ja/ and /ja.ma/ sequences (H: high and L: low) and the tonal contexts (compatible and conflicting) (Xu, 1993, 1994; He, 2010). The dependent variables were the F0 values at onset, one quarter, one half, and three quarters, and at the offset of the initial syllable in the /ma.ja/ and /ja.ma/ sequences. The following table shows the main effects and the two-way interactions of ANOVA test. Due to the phonological tone sandhi that occurs when Tone 3 is followed by another Tone 3 (dipping) (See Figure 5 (c)), the sequence Tone 3 and Tone 3 is not included in the analysis. From the table, we can see that the offset F0 value of the initial syllable significantly influence the onset F0 value at the onset, one quarter, one half, and the offset of the tone in the final syllable (target). The two-way interaction between the final tone and the offset of the initial tone was significant at the onset position of the final syllable (See Table 10). From the table we can see that the F0 value at the onset of the final syllable was subject to the influence from the offset of the initial tone ($F(1, 7) = 216.80$, $p<0.001$) and tonal contexts ($F(1, 7) = 10.09$, $p<0.05$) significantly. There was significant interaction between the tone at the onset position in the final syllable and the offset of the initial tone ($F(2, 7) = 15.12$ ($p<0.01$)), as the low initial offset significantly lowers the tone of the onset in the final syllable.

Table 10: ANOVA results for the effects of F0 offset value of the tone in the initial syllable position to the F0 tone value at five positions in the final syllable of the sequence /ma.ja/ and /ja.ma/ sequences.

position						
		0	0.25	0.5	0.75	1
Effect	df	F	F	F	F	F
Final tone (target)	3,7	9.70 (p<0.01)	4.63 (p<0.05)	152.86 (p<0.001)	75.12 (p<0.001)	299.18 (p<0.001)
Offset of the Initial tone (non_target)	1,7	216.80 (p<0.001)	6.39 (p<0.05)	10.77 (p<0.05)	2.19 (p=0.18)	24. 22 (p<0.01)
Tonal contexts	1,7	10.09 (p<0.05)	0.90 (p=0.37)	0.20 (p=0.66)	0.36 (p=0.56)	0.22 (p=0.65)
Final tone * Offset of the Initial tone	2,7	15.12 (p<0.01)	1.13 (p=0.40)	0.61 (p=0.63)	1.81 (p=0.23)	0.83 (p=0.52)

In terms of duration, a two-way ANOVA was performed on the duration of the final syllable, with four Mandarin tones in the initial syllable and the four Mandarin tones in the final syllable as independent variables. The durational variation due to the tone of the final syllable was significant ($F(3,9)=8.50$, $p<0.01$), but the variation due to the tone of the initial syllable was not. Futher analysis showed that the duration of Tone 4 (HL) in the final syllable was significantly shorter. The mean durations of the final syllables of the four tones were listed in the following table. Compared with the duration of the tones in monosyllable, the mean durations of the tones in the final syllables was shorter in general, however longer than the duration of the tones in the initial syllable, which reflects the final syllable lengthening effect, proven by several previous studies (Oller & Smith, 1977; Byrd et al., 2005). Similar with the duration rank order in initial syllables: Tone 2 (LH) was the longest tone in final position and Tone 4 (HL) was the shortest (see Table 11).

Table 11: Duration of the four Mandarin tones in the final syllable

Tone 1 (H)	Tone 2 (LH)	Tone 3 (dipping)	Tone 4 (HL)
477 ms	499 ms	457 ms	414 ms

3.5.2.3 Compatible and conflicting contexts

Spearman rank correlation (Baayen, 2008) was performed to examine the correlation between the average F0 value of the tones in monosyllable and tones in either initial or final syllable separately. It was expected that if the tone in the initial or final syllable maintained their cannoical form, then the F0 value of the tone was supposed to be correlated to the F0 value of the tone in monosyllable. That is to say, the rank correlation was expected to be stronger and statistically significant in the compatible contexts than in the conflicting context. The correlation coefficient (rho) showed the strengh and the direction of the correlation between the two variables.

Table 12 shows the the correlation between the average F0 value of the four tones in the initial syllable of 16 combinations and the average F0 value of the four tones in monosyllables by performing Spearman rank correlation test. F0 value in the initial syllable was expected to be correlated with the related F0 value in the monosyllables. To examine the tonal context effect, a two-way ANOVA test was conducted. The independent variables were the four Mandarin tones in the initial syllable and the contexts (compatible and conflicting). The dependent variables were the correlation strength (rho) between the F0 value in the initial syllable and the related F0 value in the monosyllable, obtained by performing the Spearman rank correlation test. It was found that the tone of the initial syllable effect was significant ($F(3,11)= 4.89$, $p< 0.05$), especially for the F0 value of Tone 4 (HL) in the initial syllables. However the context effect was not significant, which indicates that the F0 value of the initial syllable is not subject to the influence from the tonal contexts of compatible and conflicting contexts. In addition, it was shown that Tone 1 (H) in the initial syllable was the least tone to be correlated with any other tones in monosyllable. There was significant positive correlation between F0 value of Tone 2 (LH) in the tone sequence of two and the F0 value of Tone 2 (LH) and Tone 3 (L) in monosyllables. F0 value of Tone 3 (L) in the tone sequence of two was significantly and negatively correlated with the F0 value of Tone 2 (LH) and significantly and positvely correlated with Tone 4 (HL) in monosyllables. In addition, there was significant negative correlation between F0 value of Tone 4 (HL) in the sequence of two and the F0 value of Tone 2 (LH) in monosyllables. Also, there was significant positive correlation between F0 value of Tone 4 (HL) in the sequnece of two and the F0 value of Tone 4 (HL) in monosyllables. Therefore, according to the correlation pattern, it was expected that the confusion in the initial syllable between Tone 2 (LH) and Tone 3 (L), Tone 3 (L) and Tone 4 (HL), as well as Tone 2 (LH) and Tone 4 (HL) should be comparatively higher than other pairs for L2 learners in the following tone identification tasks in Study II.

Table 12: Correlation matrix between the tones in the initial syllables and the tones in the monosyllables

monosyllable / initial tone in the tone combination	Tone 1 (H)	Tone 2 (LH)	Tone 3 (L)	Tone 4 (HL)
Tone 1 in Tone 1-Tone 1	0.07	-0.59	0.26	0.69 ($p<0.05$)
Tone 1 in Tone 1-Tone 2	0.79 ($p<0.001$)	0.27	-0.11	-0.35
Tone 1 in Tone 1-Tone 3	0.17	0.33	-0.46	-0.45
Tone 1 in Tone 1-Tone 4	0.52	-0.04	0.28	0.07
Tone 2 in Tone 2-Tone 1	0.05	0.68 ($p<0.05$)	0.91 ($p<0.001$)	-0.51

Tone 2 in Tone 2-Tone 2	0.31	0.71 ($p<0.01$)	0.84 ($p<0.001$)	-0.59
Tone 2 in Tone 2-Tone 3	0.20	0.73 ($p<0.01$)	0.87 ($p<0.001$)	-0.59
Tone 2 in Tone 2-Tone 4	0.52	0.37	0.77 ($p<0.001$)	-0.25
Tone 3 in Tone 3-Tone 1	-0.41	-0.90 ($p<0.001$)	-0.12	0.96 ($p<0.001$)
Tone 3 in Tone 3-Tone 2	-0.42	-0.91 ($p<0.001$)	-0.12	0.97 ($p<0.001$)
Tone 3 in Tone 3-Tone 3	0.30	0.62	0.88 ($p<0.001$)	-0.47
Tone 3 in Tone 3-Tone 4	-0.19	-0.99 ($p<0.001$)	-0.37	0.96 ($p<0.001$)
Tone 4 in Tone 4- Tone1	-0.31	-0.96 ($p<0.001$)	-0.19	0.99 ($p<0.001$)
Tone 4 in Tone 4-Tone 2	-0.32	-0.96 ($p<0.001$)	-0.19	0.98 ($p<0.001$)
Tone 4 in Tone 4-Tone 3	-0.13	-0.98 ($p<0.001$)	-0.46	0.92 ($p<0.001$)
Tone 4 in Tone 4-Tone 4	-0.27	-0.98 ($p<0.001$)	-0.27	0.98 ($p<0.001$)

Note: Probability values adjusted for multiple tests.

Table 13 shows the the correlation between the average F0 value of the four tones in the final syllable of 16 combinations and the average F0 value of the four tones in monosyllables by performing Spearman rank correlation test. Similiarly, the tone in the final syllable was expected to be correlated with the related tone in monosyllables. To examine the tonal context effect, a two-way ANOVA test was conducted. The independent variables were the four Mandarin tone in the final syllable, and the tonal contexts (compatible and conflicting). The dependent variables were the correlation strength (rho) between the F0 value in the final syllable and the related F0 value in the monosyllable, obtained by performing the Spearman rank correlation test. It was found that the correlation strengh (rho) in the compatible context was significantly higher than the conflicting context ($F(1,11)= 7.62$, $p< 0.05$), and the tone effect was also significnat ($F(3,11)= 3.60$, $p< 0.05$). This indicates that the F0 value of tones in the final syllable are most subject to the influence of the tonal contexts. In addition, it was shown that there was significant positive correlation between F0 value of Tone 1 (H) in the tone sequence of two and the F0 value of Tone 2 (LH). Also, there was significant negative correlation between F0 value of Tone 1 in the tone sequence of two and the F0 value of Tone 4 (HL) in monosyllables. F0 value of Tone 2 (LH) in the tone sequence of two was significantly and positively correlated with the F0 value of Tone 3 (L) in monosyllables. Surprisingly, F0 value of Tone 3 (dipping) in the final syllable of tone sequence of two was not significantly correlated with other tones. Also, there was significant negative correlation between F0 value of Tone 4 (HL) in the sequence of two and the F0 value of Tone 2 (LH) in monosyllables and significant positive correlation between F0 value of Tone 4 (HL) in the sequence of two and the F0 value of Tone 4 (HL) in monosyllables. Therefore, according to the correlation pattern, it was expected that the confusion in the final syllable between Tone 1 (H) and Tone 2 (LH), Tone 1 (H) and Tone 4 (HL), as well as Tone 2 (LH) and Tone 4 (HL) should be comparatively higher than other pairs for L2 learners in the following tone identification tasks in Study II. Compared with the correlation patterns in the initial syllable, there was less correlation that may bring less challenge in tone identification in final syllable for L2 learners.

Table 13: Correlation matrix between the tones in the final syllables and the tones in the monosyllables

Note: Probability values adjusted for multiple tests.

monosyllable / final tone in the tone combination	Tone 1 (H)	Tone 2 (LH)	Tone 3 (L)	Tone 4 (HL)
Tone 1 in Tone 1-Tone 1	0.37	-0.08	0.63	0
Tone 1 in Tone 2-Tone 1	0.88 (p< 0.001)	0.16	0.12	-0.22
Tone 1 in Tone 3-Tone 1	0.49	0.84 (p< 0.001)	0.02	-0.91 (p< 0.001)
Tone 1 in Tone 4-Tone 1	0.59	0.79 (p< 0.001)	0.03	-0.88 (p< 0.001)
Tone 2 in Tone 1-Tone 2	0.25	0.26	0.87 (p<0.001)	-0.10

Tone 2 in Tone 2-Tone 2	0.20	0.16	0.89 (p<0.001)	0.01
Tone 2 in Tone 3-Tone 2	0.44	0.89 (p<0.001)	0.18	-0.94 (p<0.001)
Tone 2 in Tone 4-Tone 2	0.55	0.72 (p<0.01)	0.34	-0.76 (p<0.001)
Tone 3 in Tone 1-Tone 3	0.14	-0.63	0.07	0.63
Tone 3 in Tone 2-Tone 3	0.14	-0.57	0.27	0.63
Tone 3 in Tone 3-Tone 3	0.07	-0.18	0.68 0 (p< 0.05)	0.35
Tone 3 in Tone 4-Tone 3	0.02	-0.19	0.62	0.30
Tone 4 in Tone 1-Tone 4	-0.41	-0.91 (p<0.001)	-0.13	0.97 (p<0.001)
Tone 4 in Tone 2-Tone 4	-0.36	-0.94 (p<0.001)	-0.14	0.98 (p<0.001)
Tone 4 in Tone 3-Tone 4	-0.14	-0.94 (p<0.001)	-0.65 (p< 0.05)	0.85 (p<0.001)
Tone 4 in Tone 4-Tone 4	-0.16	-0.96 (p<0.001)	-0.56	0.89 (p<0.001)

3.6 Discussion

The present study offered a combined view to systematically examine tonal co-articulation concerning tone sequence of disyllabic words in Mandarin. It has been hypothesized that the F0 height and F0 contour of the tone sequence of two deviates from the tone in monosyllable Mandarin lexical tones, as the placement of the tone in either the initial or final syllable position, or the compatible and conflicting tonal context would influence the F0 height and F0 contour.

It has been found that the F0 patterns of tones produced in isolation can be described the same as the F0 patterns when produced in context, which can be described as Tone 1 (H), Tone 2 (LH), Tone 3 (L), and Tone 4 (HL) (Shih, 1987; Xu, 1997). The F0 correlation matrix of tones in monosyllables shows that the rankings of F0 value of Tone 2 (LH) and Tone 4 (HL) was significantly and negatively correlated, and there was moderate correlation between Tone 2 (LH) and Tone 3 (L).

F0 contour and F0 height of the tone in the initial and final syllable was different from the tone in monosyllables. For the effect of the syllable position, the F0 contour and F0 height at the onset of final syllables deviated more from the canonical form in monosyllables than at the offset of the initial syllable in the nonsense sequences, which is consistent with previous studies (Xu, 1997). Specifically, the tone onset in the final syllable were heightened by the offset of the initial Tone 1 (H) and Tone 2 (LH), and lowered by offset of the initial Tone 3 (L) and Tone 4 (HL). However, the F0 value and contour of the tone in the initial syllable was less influeced by the tone onset in the final syllable. It has been observed that lower F0 value of the tone in the final syllable tend to increase the offset F0 value of tone in the initial syllable. It is confirmed that the carry over is assimilatory and the magnititude is bigger, and the anticipatory effect is dissimilatory and the magnitude is smaller (Xu, 1994; Xu, 1997).

Concerning the tonal context effect, including the compatible and conflicting contexts, the F0 value in the final syllable was signficantly subject to the influence of the tonal context. The F0 value of the tone in the initial syllable was less likely to be influenced by the tonal contexts, as it was found that the compatible and conflicting context effect was not significant in the initial syllable but significant in the final syllable. If L2 learners give their primary cue to the onset F0 value and contour of the tone and variation magnitude matter, then the final syllable may bring more difficulty in non-native tone identification than the initial syllable. It is expected the high conflusion pattern in the final syllable should include the following tone pairs: Tone 1 (H) and Tone 2 (LH), Tone 1 (H) and Tone 4 (HL), Tone 2 (LH) and Tone 3 (L), as well as Tone 2 (LH) and Tone 4 (HL). However, if L2 learner give their primary perceptual cue to the offset of the F0 value and contour of the tone, and the magnitude is less relevant, then the tone in the initial syllable can also be more challenging for L2 learners to identify tones in context. In addition, F0 values in the initial syllable were less correlated to the related F0 value in monosyllables, which may also bring more challenge to the tone identification tasks of the initial syllable for L2 learners. It is expected that the high conflusion pattern, which may appear in the initial syllable, includes tone pairs: Tone 2 (LH) and Tone 3 (L), Tone 3 (L) and Tone 4 (HL), as well as Tone 2 (LH) and Tone 4 (HL). In addition, F0 value of Tone 4 (HL) in the both initial and final syllables was found to be more correlated with the F0 value of Tone 4 (HL) in the monosyllables and less subject to the tonal contexts influence. This indicates that Tone 4 (HL) might be easier for L2 learners to identify in the following tone identification tasks than other tones.

In summary, this study revealed how F0 contour and F0 height in the production of L1 Chinese can be influenced by both tonal combinations and placements in nonsense sequences, which may benefit teachers in predicting the points of difficulty in learning Mandarin tones and in helping students improve their tone identification and production.

It would be interesting to explore how the syllable position, and tonal contexts (compatible and conflicting contexts) will influence the nonnative tone identification. In addition, whether the variation of the tone onset in the final syllable and the tone offset in the initial syllable may increase the difficulty in tone identification tasks. Whether the acoustic variation in the stimuli and the correlation matrix would predict the L2 identification accuracy and error patterns. Therefore, in the following chapter, three tasks concerning the identification of mandarin tones in context has been conducted to examine how the aforementioned acoustic variation and correlation would influence the identification of Mandarin lexical tones of both native Mandarin-speaking listners and L2 learners of Chinese.

CHAPTER 4 STUDY II

THE IDENTIFICATION OF MANDARIN LEXICAL TONES IN CONTEXT

The current study tested the identification of Mandarin tones by L1 English learners of Chinese in both monosyllables and disyllabic non-words by using DMDX (Forster, K. I., & Forster, J. C., 2003); in terms of how accuracy rates, identification sensitivity, error patterns, and reaction times are influenced by tone, syllable position, tonal context, and learning experience. The experimental design employed tone identification tasks, including three experimental conditions: monosyllable, the initial syllable and the final syllable of disyllabic non-words. There were a total of 46 participants, including three experimental groups and a control group of native Chinese subjects, who took identical tests, and were included to serve as a baseline.

4.1 Research Questions

Generally, the present experiment was designed to test how accuracy rates, identification sensitivity, error patterns, and reaction times are influenced by tone, syllable position, tonal context, and learning experience.

Specifically: The current study was based on the following research questions:

- What are the similarities and differences, including the accuracy rates, identification sensitivity, error patterns, and reaction times, due to tone and syllable position among tones in monosyllables, the initial syllable of a disyllable, and the final syllable of the disyllable?
- How do compatible and conflicting contexts influence the accuracy rates and reaction time?

- Does L1 English learners of Chinese' identification of Mandarin two tone sequences improve with Mandarin learning experience? What is the difference between native Mandarin speakers and L1 English learners of Chinese?

4.2 Hypotheses and Expected Results

Five hypotheses were tested throughout the present study.

- The tones in monosyllables will be identified with the highest accuracy and sensitivity, and shortest reaction time, followed by tones in disyllables.
- Syllable position will have significant effects on accuracy and reaction time. If L2 learners give their primary cue to the onset F0 value and contour of the tone and variation magnitude matter, then the final syllable may bring more difficulty in non-native tone identification than the initial syllable, with lower accuracy and longer reaction time. However, if L2 learners give their primary perceptual cue to the offset of the F0 value and contour of the tone, and the magnitude is less relevant, then the tone in the initial syllable can be more challenging for L2 learners to identify, with lower accuracy and longer reaction time.
- Higher accuracy and faster reaction time is expected in the compatible context than the conflicting contexts.
- The confusion between Tone 2 (LH) and Tone 3 (L) will be most salient, among the three experimental conditions, including monosyllable, the initial syllable and the final syllable of disyllabic non-words. Tone 4 (HL) is expected to be the easiest tone for all three experimental conditions. In terms of error patterns in context, the expected error patterns of tones from the initial syllable in the disyllables include the following tone pairs: Tone

2 (LH) and Tone 3 (L), Tone 3 (L) and Tone 4 (HL), as well as Tone 2 (LH) and Tone 4 (HL). The expected error patterns of tones from final syllable in the disyllables include: Tone 1 (H) and Tone 2 (LH), Tone 1 (H) and Tone 4 (HL), Tone 2 (LH) and Tone 3 (L), as well as Tone 2 (LH) and Tone 4 (HL).

- Tone 2 (LH) and Tone 3 (L) are expected to be with lower sensitivy than Tone 4 (HL) and Tone 1 (H).
- L1 English learners of Chinese' perception of Mandarin two tone sequence will improve with more Mandarin learning experience. In terms of the overall result of the Mandarin lexical tone identification, L2 learners with more learning experience will be more resemblant to those of native speakers of Mandarin, and therefore differ from the results of learners of less learning experinece.

4.3 Methods

4.3.1 Participants

Three experimental groups and one control group participated in the present study (See Table 14). In total, there were 46 participants, with no special training in music. All were students between the ages of 20 and 30 at Indiana University Bloomington.

Table 14: Participants

Group	Group	Number	L1	Years of learning
Experimental groups	1	11 (6 Female)	English	1
	2	11 (5 Female)	English	2
	3	12 (5 Female)	English	3
Control group	NS	12 (7 Female)	Mandarin Chinese	NA

The three experimental groups (Group 1, Group 2 and Group 3) consisted of L1 English learners of Chinese. Thirty-four participants' results were selected and analyzed in the current study (Four participants were excluded due to interfering heritage language experience and one participant was excluded as the testing result was substantially different from other group members).

There were 11 participants in Group 1 (6 female), who had completed first year Chinese, which means they received 150 hours of formal instruction time (15 weeks per semester, 5 hours per week). All of the participants from Group 1 were without any previous experience in learning Chinese before taking first year Chinese. None of the participants has been to a Chinese-speaking country.

There were 12 participants in Group 2 (5 female), who had completed second year Chinese, which means they received 300 hours of formal instruction time (15 weeks per semester× 4 semesters, 5 hours per week). Nine out of eleven participants finished the second level of Chinese in two years with a four-month summer break in between. Two participants finished second year Chinese by attending the Flagship Chinese Institute summer intensive program and two regular semesters. Two participants have been to a Chinese-speaking country for less than a year.

There were twelve participants in Group 3 (5 female), who had completed third year Chinese, which means they received 450 hours of formal instruction time (15 weeks per semester× 6 semesters, 5 hours per week). Seven out of twelve participants finished third year Chinese by attending the Flagship Chinese Institute summer intensive program and four regular semesters continuously. Two participants have been to a Chinese-speaking country for two years.

The control group consisted of 12 native speakers of Chinese (7 female), who underwent the same testing, and were included to serve as a baseline. Nine out of twelve are from Beijing or the northern part of China, as Beijing dialect has the most similar tonal system to that of Mandarin Chinese. The average age of the control group is 25.

All methods reported herein were approved by the Institutional Review Board at the Indiana University-Bloomington and all participants provided written informed consent.

4.3.2 Stimuli

The stimuli used in the present study were obtained from the production of native Mandarin speakers in Study I. Stimuli for identification includes 32 tokens (4 tones × 2 syllables × 2 informants × 2 repetitions) of monosyllables and 128 tokens (4 initial tones × 4 final tones × 2 syllables × 2 informants × 2 repetitions) of non-word disyllables. See Table 3 and Table 4 for the stimuli. All the stimuli amplitude were normalized with the same output of 75 dB and programed using a DMDX package (Forster, K. I., & Forster, J. C., 2003) with four input answer keys.

Twelve practice stimuli were recorded by a male speaker from Beijing, which were used to familiarize participants with the experimental conditions (See Appendix II).

4.3.3 Procedure

The experimental design employed tone identification tasks. In this design, the tone identification of L1 English learners of Chinese was examined through three experimental conditions depending on the syllable that carries the identification target, including monosyllable (Task 1), the initial syllable of disyllabic non-words (Task 2), and the final syllable of disyllabic non words (Task 3) (See Table 15). In Task 1, participants heard one monosyllable and the identification target was the tone of the syllable for each trail. In Task 2, participants heard one disyllabic non-word and the identification target was the tone of the initial syllable of the disyllabic non-word for each trial. In Task 3, participants heard one disyllabic non word and the identification target was the tone of the final syllable of the disyllabic non-word for each trial. All participants in the experimental groups and control group heard 160 stimuli. The order of the three tasks and the trial order in each task were randomized for every participant.

Table 15: Tasks

Tasks	Tokens	No. of syllable heard for each trial	Identification target
Task 1: Monosyllable	32	1	The tone in monosyllable
Task 2: The initial syllable of disyllabic non words	64	2	The tone in the initial syllable
Task 3: The final syllable of disyllabic non words	64	2	The tone in the final syllable
Total	160		

DMDX (Forster, K. I., & Forster, J. C., 2003) was used to run the experiment. See Figure 6 for the experimental procedure.

The participants first were seated in a quiet room, and then asked to listen to the stimuli and complete the tone identification test individually using headphones. Within each experiment, participants first completed the familiarizing trials, which included four practice trials. Although no feedback was given, any of their questions about the experimental procedure were answered. After being familiarized with the experimental condition, they started the real experimental task. Instructions for both the familiarizing task and experimental task were provided (See Appendix III).

Then at the beginning of each trial, four diacritics of Mandarin lexical tones “¯ / v \” were displayed on the center of the computer screen. It is a common practice to transcribe Mandarin lexical tones with diacritics “¯ / v \” over the vowels. Also, it is mandatory for all the participants to associate the four diacritics with the category of the Mandarin lexical tones by the very first month of learning Mandarin Chinese. Therefore, L1 English learners of Chinese are very familiar with these four diacritics, which visually symbolize the F0 height and F0 contour of Mandarin lexical tones, including Tone 1 (H) “ ¯ ”, Tone 2 (LH) “/”, Tone 3 (L /dipping) “v”, and Tone 4 (HL) “\”.

Then the participants heard the stimuli in each trial that carried the identification target. They were instructed to respond only to the tone of the target syllable. Participants used the keyboard as a response input method to answer the tone of the target syllable they just heard. All the participants were asked to enter their response as soon as possible. Their reaction time was recorded from the onset of the target syllable to the time of response. However, if the participants took too much time, beyond the sum of syllable duration, which was the time limit (2000ms) to give their response, their response was considered as void and it moved on to the next token for the identification task automatically.

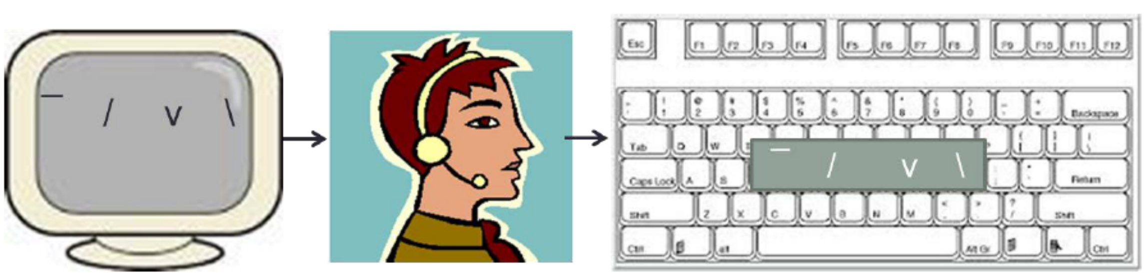


Figure 6: Experimental procedure

The tone identification task took about 20 minutes to finish. Including the language background questionnaire (See Appendix IV), briefing, instructions, practice trials and the actual experiments, each subject spent no more than 30 minutes completing this task.

4.4 Data Analysis

Within each task, analyses were conducted on accuracy, sensitivity index (d' or d prime), error patterns, and reaction time of tone identification responses.

For the Mandarin tone identification task, a higher score (the proportion of correct responses) corresponds to better identification ability. For each trial, the accuracy was binary coded with 0 (incorrect) and 1 (correct). Based on the binomial distribution of the identification scores, the data was further being transformed in a format in which the number of correct responses and incorrect responses for each line in the data frame, and thereby to use a generalized linear model (e.g. logistic regression) (Baayen, 2008; Xu et al., 2006; Ning, submitted) as an appropriate approach. For the present study, in each group, the subjects were randomly selected. Therefore, subjects were included in the present study as a random effect, assuming that individual specific effects are uncorrelated with the independent variables (Littell et al., 2002). To include the fixed effects (target tone, syllable position, tonal contexts, which includes compatible and conflicting contexts, and group), and the random effect (subjects), the present study used a generalized linear mixed model (GLMM) from the package 'lme4' (Bates et al., 2014) for R (R Development Core Team, 2008), an extension of generalized linear models, which allows generalization of the inferences beyond the sample used in the model (Littell et al., 2002). The R package 'multcomp' (Hothorn et al., 2014) was used to conduct pairwise comparisons for group effect by using Tukey HSD (honest significant difference).

The sensitivity index (d' or d prime) was calculated according to signal detection theory (Macmillan & Creelman, 2005; McGuire, 2010). According to signal detection theory, the sensitivity index (d' or d prime) is the z-transformation of the hit rate (H), the proportion of correct responses, minus the false-alarm rate (FA), the proportion of incorrect responses.

$$d' = z(H) - z(FA)$$

The sensitivity index (d' or d prime) has been applied to calculate the sensitivity in Mandarin lexical tone identification tasks (Chen & Massaro, 2008; Krenmayr et al., 2011). Index d prime indicates how well the participant distinguishes one lexical tone from the others. The bigger the d prime value, the better the participant is at recognizing the tone (Chen & Massaro, 2008). Specifically, according to the method used in (Krenmayr et al., 2011), the hit rate of a target tone was calculated by the number of the correct responses of the target tone (e.g. Tone 1) divided by the number of the total trials containing that target tone (e.g. Tone 1). The false alarm rate of a target tone was estimated by dividing the number of incorrect responses of the target tone (e.g. Tone 1) by the cumulative number of the other three Mandarin lexical tone trials (e.g. Tone 2, Tone 3 and Tone 4). Log-linear correction, a procedure of adding 0.001 to all response counts for calculating hit rate and deducting 0.001 to all the response counts for false alarm, were performed to the present data (as cited in Krenmayr et al., 2011). This is to avoid proportions equalizing zero or unity, which would leave the d prime value undefined, as the z-scores are infinite in these cases. Generally, the upper limit in the data is determined by the number of presentations and the actual correct responses for the ideal proportions. The negative values indicate that a subject consistently confuses the target tone with other Mandarin lexical tones. A d prime value of 0 indicates that a subject has no ability to discriminate the target tone with the other three Mandarin lexical tones. The R package 'lme4' (Bates et al., 2014) was used to construct linear mixed effects models (LMMs) and the R package 'lmerTest' (Kuznetsova et al., 2014) was used to conduct Analyses of variance (ANOVAs) on sensitivity index (d prime), with the main effects (target tone, syllable position, tonal context (compatible and conflicting contexts), and group), and random effect (subjects), by using type-III ANOVAs with denominator degrees of freedom calculated based on Satterthwaite's approximation.

The error patterns were presented with the frequency and the percentages. Twelve possible error patterns were analyzed by using two-way ANOVAs with the dependent variable: error rate, and the main effects: group and 12 error patterns, including Tone 1-to-Tone 2 (which indicates the target syllable Tone 1 being identified as Tone 2), Tone 1-to-Tone 3, Tone 1-to-Tone 4, …, and Tone 4-to-Tone 3, which has been used by a previous study (Hao, 2012). For each individual error pattern, t statistics were reported.

Reaction time was measured from stimulus offset to avoid the potential confounding of stimulus duration differences, which have been used by a previous study (Lee et al. 2010 b). Only correct responses were included in the reaction time analysis. The main effects include target tone, syllable position, tonal context, and experience. Analyses of variance (ANOVAs) were conducted on reaction time with the main effects of target tone, syllable position, tonal context (compatible and conflicting contexts), and group.

Please see the following table for the coding of the independent variables. Groups were coded with Group NS (native speaker), Group 1 (with one year Chinese learning experience), Group 2 (with two years Chinese learning experience) and Group 3 (with three years Chinese learning experience). Target tone was coded with Tone 1(H), Tone 2 (LH), Tone 3(L), and Tone 4 (HL). Syllable position was coded with monosyllable, initial syllable (in non-word disyllables), and final syllable (in non-word disyllables). Also, only for the two experimental conditions, including initial or final syllable in non-word disyllables, there was one more independent variable: compatible and conflicting contexts.

Table 16: Coding for independent variables

Fixed effects	Target Tone	Tone 1,Tone 2,Tone 3,Tone 4
	Syllable Position	monosyllable initial syllable final syllable
	Compatible and Conflicting	Compatible Conflicting
	Group	Group 1, Group 2, Group 3, Group NS
Random effects	Subjects	s1~s46

4.5 Results

4.5.1 Task 1: Monosyllable

4.5.1.1 Accuracy Rate

Figure 7 shows the four groups' accuracy of the four Mandarin lexical tones in monosyllables. Each line represents one group, including Group 1, Group 2, Group 3, and Group NS. The X-axis indicates the identification targets, including Tone 1 (H), Tone 2 (LH), Tone 3 (L) and Tone 4 (HL), and the Y-axis indicates the accuracy mean, the proportion of correct responses.

According to the descriptive statistics, the accuracy mean of Group NS was the highest compared with the other three learner groups, including Group 1, Group 2 and Group 3. The accuracy means of different groups from high to low were Group NS (0.990) > Group 3 (0.883) > Group 2 (0.844) > Group 1 (0.821). The accuracy means of learner groups regarding tones in monosyllables from high to low were Tone 1 (0.937) > Tone 4 (0.870) > Tone 2 (0.833) > Tone 3 (0.758). Compared with learner groups, the accuracy means of Group NS of tones in monosyllables from high to low were Tone 1 (1.000) = Tone 4 (1.000) > Tone 3 (0.990) > Tone 2 (0.969), which is near the unit ceiling.

A generalized linear mixed model (GLMM) was used to regress the identification accuracy for tones in monosyllables of each trial (1 for correct and 0 for incorrect) for group, tone, and the interactions between them. The full model showed that there was no significant interaction between group and tone effects (Wald χ^2 (9)) = 7.74, p = 0.56). The reduced model generated by removing the interaction term showed that there was a significant main effect of group (Wald χ^2 (3)) = 34.46, $p < 0.001$) and tone (Wald χ^2 (3)) = 40.51, $p < 0.001$). Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that the L2 learner groups did not perform statistically differently among each other, while they were all significantly different from Group NS ($p < 0.001$).

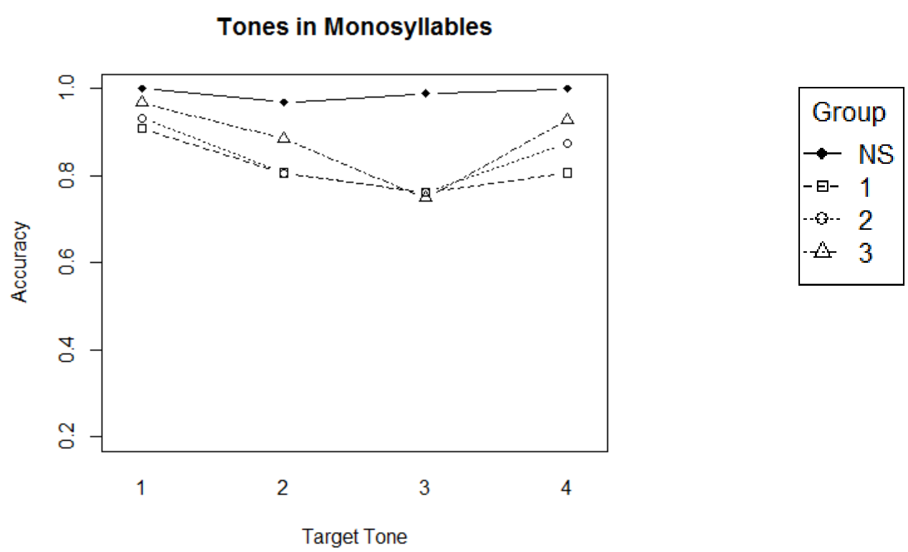


Figure 7: Accuracy of tones in monosyllables

4.5.1.2 Identification Sensitivity (d prime)

Figure 8 shows the four groups' identification sensitivity of the four Mandarin lexical tones in monosyllables.

According to the descriptive statistics, Group NS had the highest identification sensitivity compared with the other three learner groups. The identification sensitivities of different groups from high to low were Group NS (7.295) > Group 3 (4.993) > Group 2 (4.338) > Group 1 (4.233). The identification sensitivities of learner groups from high to low were Tone 1 (5.859) > Tone 4 (5.580) > Tone 3 (3.527) > Tone 2 (3.174). Therefore, Tone 1 (H) and Tone 4 (HL) were comparatively easier than Tone 2 (LH) and Tone 3 (L) for learner groups to distinguish from other tones, and Tone 2 (LH) was the most difficult for all the learner groups. For Group NS, the identification sensitivity from high to low was Tone 1 (7.596) = Tone 4 (7.596) > Tone 3 (7.203) > Tone 2 (6.785). There are mainly two differences when comparing Figure 8 with Figure 7, namely calculating identification sensitivity versus solely calculating identification accuracy. First, Tone 2 (LH), instead of Tone 3 (L) had the lowest identification sensitivity for all the learner groups. Second, with more learning experience, the identification sensitivity for Tone 3 (L) also improved considerably, which is very different from their similar identification accuracy.

A linear mixed model (LMM) was used to regress the identification sensitivity in monosyllable of each trial for group, tone, and the interactions between them. The full model showed that there was a significant main effect of group ($F (3, 42) = 13.898$, $p < 0.001$) and tone ($F (3, 126) = 45.935$, $p < 0.001$), as well as significant interaction between group and tone effects ($F (9, 126) = 2.803$, $p < 0.001$). Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that L2 learner groups' identification sensitivity did not significantly differ among each other, while they were all significantly different from Group NS (p <0.001).

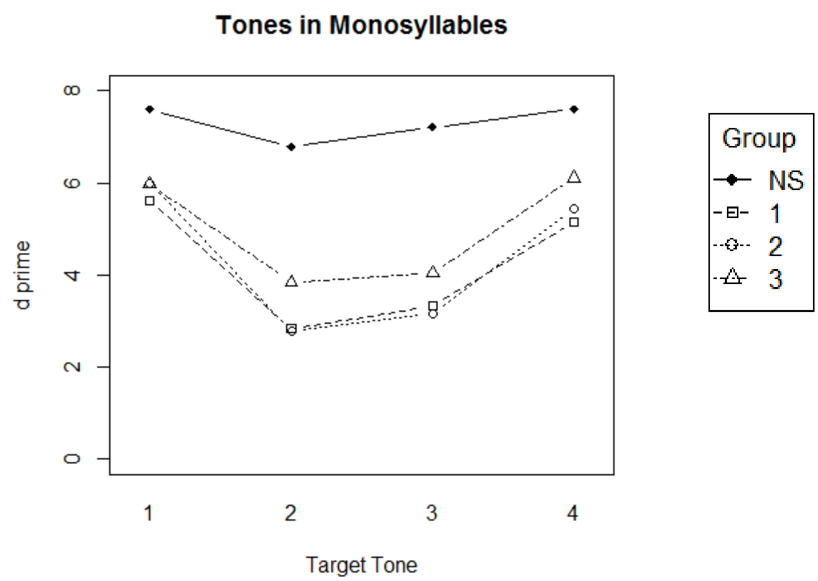


Figure 8: Identification sensitivity of tones in monosyllables

4.5.1.3 Error Pattern

Table 17 shows the error patterns of Mandarin four tones in monosyllable with the percentages in parentheses.

Two-way ANOVA results showed that the main effect of group ($F(3,33)= 5.928$, $p<0.01$) and error pattern ($F(11,33)= 7.924$, $p<0.001$) was significant. Group 1 ($t(33)=3.853$, $p<0.001$), Group 2 ($t(33)=3.409$, $p<0.01$), and Group 3 ($t(33)=2.487$, $p<0.05$) were all significant with more errors than Group NS. Generally, among the learner groups, with more learning experience there were less errors in tone identification. Tone 3 (L) was more frequently incorrectly identified as Tone 2 (LH) than the correspondent Tone 2 (LH) was incorrectly identified as Tone 3 (L) among learner groups. Tone 3 (L)-to-Tone 2 (LH) (which indicates the target syllable Tone 3 (L) being identified as Tone 2 (LH)) was significantly higher than other error patterns (t (33)

=6.306, p<0.001), which indicates that to distinguish Tone 3 (L) from Tone 2 (LH) is the most difficult pairing.

Table 17: Confusion matrix for tones in monosyllables with the percentages in parenthesese

Group	Stimulus	Response				
		Tone 1 (H)	Tone 2 (LH)	Tone 3 (L)	Tone 4 (HL)	NA
Group 1	Tone 1 (H)	**80 (90.9)**	4 (4.5)	1 (1.1)	2 (2.3)	1 (1.1)
	Tone 2 (LH)	7 (8.0)	**71 (80.7)**	5 (5.7)	1 (1.1)	4 (4.5)
	Tone 3 (L)	1 (1.1)	17 (19.3)	**67 (**76.1)	1 (1.1)	2 (2.3)
	Tone 4 (HL)	3 (3.4)	6 (6.8)	5 (5.7)	**71 (80.7)**	3 (3.4)
Group 2	Tone 1 (H)	**82 (93.2)**	3 (3.4)	0 (0)	2 (2.3)	1 (1.1)
	Tone 2 (LH)	4 (4.5)	**71 (80.7)**	9 (10.2)	0 (0)	4 (4.5)
	Tone 3 (L)	1 (1.1)	19 (21.6)	**67 (76.1)**	0 (0)	1 (1.1)
	Tone 4 (HL)	0 (0)	7 (8.0)	2 (2.3)	**77 (87.5)**	2 (2.3)
Group 3	Tone 1 (H)	**93 (96.9)**	1 (1.0)	0 (0)	1 (1.0)	1 (1.0)
	Tone 2 (LH)	4 (4.2)	**85 (88.5)**	4 (4.2)	0 (0)	3 (3.1)
	Tone 3 (L)	0 (0)	21 (21.9)	**72 (75.0)**	1 (1.0)	2 (2.1)
	Tone 4 (HL)	4 (4.2)	1 (1.0)	1 (1.0)	**89 (92.7)**	1 (1.0)
Group NS	Tone 1 (H)	**96 (100)**	0	0	0	0
	Tone 2 (LH)	0	**93 (96.9)**	1 (1.0)	0	2 (2.3)
	Tone 3 (L)	0	1 (1.0)	**95 (99.0)**	0	0
	Tone 4 (HL)	0	0	0	**96 (100)**	0

4.5.1.4 Reaction Time

According to the descriptive statistics, the reaction time for identification of the four Mandarin lexical tones in monosyllables, Group NS had the least reaction time compared with the other three learner groups. The reaction time of the four Mandarin lexical tones of different groups from short to long in terms of miliseconds were Group NS (473.584) < Group 2 (612.687) < Group 3 (633.365) < Group 1 (660.475). For Group NS, the reaction time from short to long were Tone 3 (406.983) < Tone 1 (458.934) < Tone 2 (496.210) < Tone 4 (532.224). For learner groups, Tone 1 (532.986) < Tone 3 (592.744) < Tone 4 (692.034) < Tone 2 (729.196). Generally, Tone 1 (H) and Tone 3 (L/dipping) had a shorter reaction time to identify than Tone 2 (LH) and Tone 4 (HL) for both Group NS and learner groups, as shown in Figure 9.

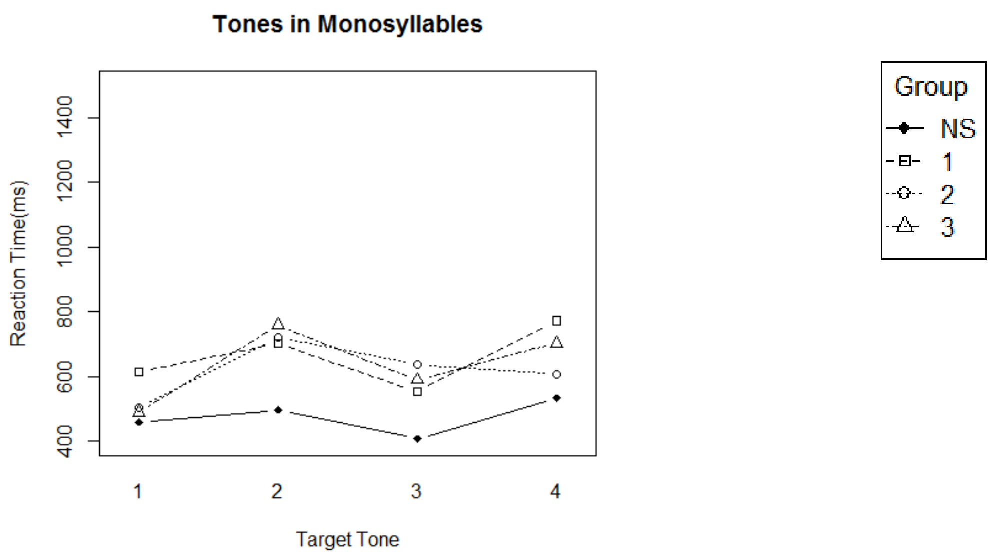


Figure 9: Reaction time of tones in monosyllables

A linear mixed model (LMM) was used to regress the identification reaction time for monosyllables of each trial for group, tone, and the interactions between them. The full model showed that there was a significant main effect of group ($F(3, 42) = 4.355$, $p < 0.01$) and tone ($F(3, 1248) = 23.959$, $p < 0.001$), as well as significant interaction between group and tone effects

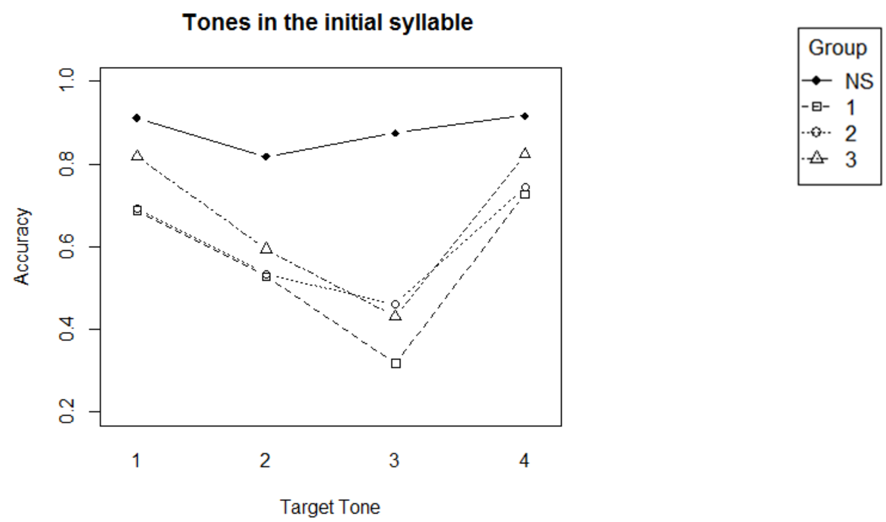
(F (9, 1248) = 3.690, $p < 0.001$). Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that the L2 learner groups' identification reaction time was not significantly differently among each other, while Group 1 and Group 3 were both significantly different from Group NS ($p < 0.05$).

To summarize, learner groups' accuracy and identification sensitivity was significantly lower and their identification reaction time was significantly slower than native Mandarin speakers. L2 Learners with more learning experience were more resemblant to those of native speakers of Mandarin, though L2 learner groups did not perform statistically differently among each other. For learner groups, Tone 3 (L) was with the lowest acuracy and Tone 2 (LH) was with the lowest identification sensitivity. For Group NS, Tone 2 (LH) was with the lowest acuracy and lowest identification sensitivity. The error Tone 3 (L)-to-Tone 2 (LH) (which indicates the target syllable Tone 3 (L) being identified as Tone 2 (LH)) was the most error type. Tone 1 (H) and Tone 3 (L) took shorter reaction time to identify than Tone 2 (LH) and Tone 4 (HL) for both Group NS and learner groups.

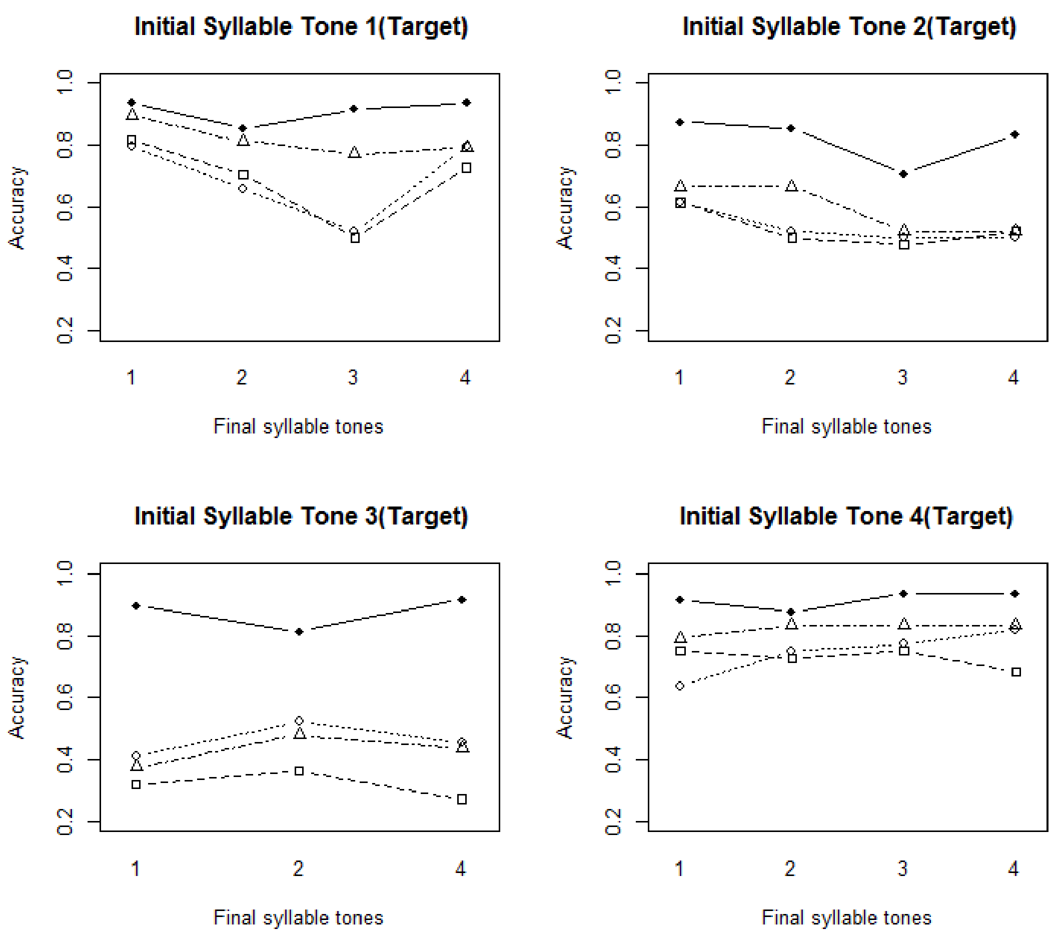
4.5.2 Task 2: The Initial Syllable of Disyllabic Non-words

4.5.2.1 Accuracy Rate

Figure 10 shows the four groups' accuracies of the four Mandarin lexical tones in the initial syllable of disyllabic non-words. The combination Tone 3-Tone 3 was excluded from the figure, due to the phonological tone sandhi that occurs when Tone 3 is followed by another Tone 3: the first Tone 3 (L) changes to the tone with F0 value LH, which is similar to Tone 2 (LH). According to the production result in Study I, there was no significant difference concerning the F0 value between the initial Tone 3 in the combination Tone 3-Tone 3 and Tone 2, therefore, the combination Tone 3-Tone 3 was separately analyzed.



(a)



(b)

Figure 10:Accuracy rate of the tones in the initial syllable

According to the descriptive statistics, the accuracy mean of Group NS was the highest compared with the other three learner groups. The accuracy mean of different groups from high to low was Group NS (0.881) > Group 3 (0.682) > Group 2 (0.618) > Group 1 (0.581). The accuracy means of learner groups regarding tones in the initial syllable from high to low are Tone 4 (0.767) > Tone 1 (0.735) > Tone 2 (0.553) > Tone 3 (0.404). We can extrapolate from the figure that Tone 1 (H) and Tone 4 (HL) were comparatively easier to identify than Tone 2 (LH) and Tone 3 (L) for learner groups, and Tone 3 (L) was the most difficult to identify for all the learner groups. More learning experience did not improve the identification of Tone 3 (L) in terms of accuracy rate. For Group NS, the accuracy means of tones in the initial syllable from high to low are Tone 4 (0.917) > Tone 1 (0.911) > Tone 3 (0.875) > Tone 2 (0.818).

GLMM was used to regress the identification accuracy for the tones in the initial syllable of each trial (1 for correct and 0 for incorrect) for group, target tones in initial syllable, the tones in final syllable and the interactions between target tones and group as well as target tones and the tones in final syllable. The model showed that there was significant main effect of group (Wald $\chi 2$ (3)) = 16.924, $p < 0.001$) and target tone (Wald $\chi 2$ (3)) = 19.674, $p < 0.001$), however the effect of tones in final syllable was not significant (Wald $\chi 2$ (1)) = 2.99, $p = 0.083$). In addition, there was significant interaction between target tones and group (Wald $\chi 2$ (9)) = 21.390, $p < 0.05$), and between target tones and the tones in final syllable (Wald $\chi 2$ (3)) = 11.390, $p < 0.05$) . Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that the L2 learner groups did not perform statistically differently among each other, while they were all significantly different from Group NS ($p <0.001$). Also, the accuracy of Tone 2 (LH) ($z = -7.884$, $p<0.001$) and Tone 3 (L) ($z = -11.032$, $p<0.001$) was significantly lower, according to the estimated regression coefficients along with Z statistics. Specifically, when the initial syllable Tone 1 (H) and Tone 4 (HL) was the target, both Group1 and Group2 were significantly different from Group NS ($P<0.05$), however no significant difference found for Group 3, which indicates with more learning experience, learners' identification of Tone 1 (H) and Tone 4 (HL) can markedly improve.

Concerning the identification accuracy of initial syllable Tone 3 in the Tone 3-Tone 3 combination, both response Tone 2 and response Tone 3 were counted as correct response, due to the phonological tone sandhi of Tone 3 (L). The accuracy means for learner groups were all around 75%, and for Group NS was above 90%. Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that there was no significant group difference between learner groups and Group NS.

To further explore how identification accuracy of the tones in the initial syllable was influenced by both tones in initial syllable and final syllable, the context effect including compatible and conflicting context on tone idenficiation accuracy of the tones in the initial syllable was included. According to the descriptive statistics, the identification accuracy was higher in compatible context than in conflicting context, including learner groups on average : compatible context (0.682) > conflicting context (0.583) and Group NS : compatible context (0.887) > conflicting context (0.875). GLMM showed that there was significant context effect (Wald $\chi 2$ (1)) = 24.905, $p < 0.001$). In addition, it was found that with additional learning experience, the effect of the compatible/conflicting context decreased (See Figure 11).

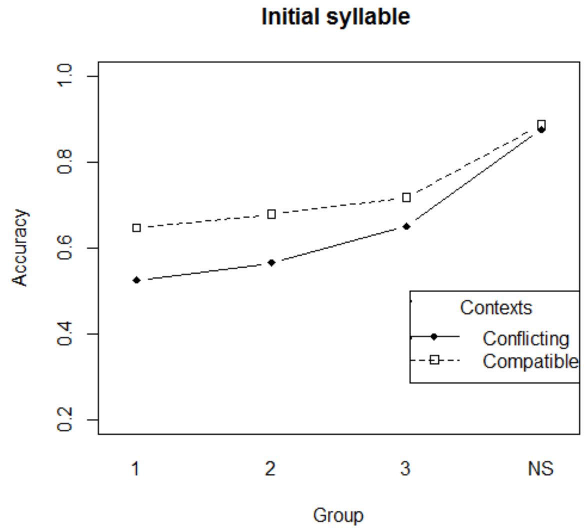


Figure 11: Context effects on accuracy in the initial syllable

4.5.2.2 Identification Sensitivity (d prime)

Figure 12 shows the four groups' identification sensitivity of the four Mandarin lexical tones in the initial syllable.

According to the descriptive statistics, the identification sensitivity of the four Mandarin lexical tones in the initial syllables, Group NS was the highest compared with the other three learner groups. The identification sensitivity of different groups from high to low was Group NS (5.547) > Group 3 (2.679) > Group 2 (2.022) > Group 1 (1.771). The identification sensitivity of learner groups regarding tones in the initial syllables from high to low was Tone 1 (2.942) > Tone 4 (2.559) > Tone 2 (1.809) > Tone 3 (1.381). For Group NS, the identification sensitivity of tones in the initial syllable from high to low was Tone 4 (6.849) > Tone 1 (6.616) > Tone 3 (4.430) > Tone 2 (4.293). When comparing Figure 12 with Figure 10 (a), namely calculating identification sensitivity versus solely calculating identification accuracy, the biggest difference is that the identification sensitivity for Tone 3 (L) and Tone 4 (HL) was associated more with learning experience, as there was no overlap for Tone 3 (L) and Tone 4 (HL) identification sensitivity. This is in contrast with the overlap for Tone 3 (L) and Tone 4 (HL) identification accuracy, which indicates that identification sensitivity has better discriminating power than identification accuracy.

LMM was used to regress the tone identification sensitivity in the initial syllable of each trial for group, tone, and the interactions between them. The full model showed that there was significant main effect of group ($F (3, 42) = 10.826, p < 0.001$) and tone ($F (3, 126) = 34.004, p < 0.001$), as well as significant interaction between group and tone effects ($F (9, 126) = 2.217, p < 0.05$). Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that the L2 learner groups' identification sensitivity was not significantly differently among each other, while they were all significantly different from Group NS ($p <0.01$).

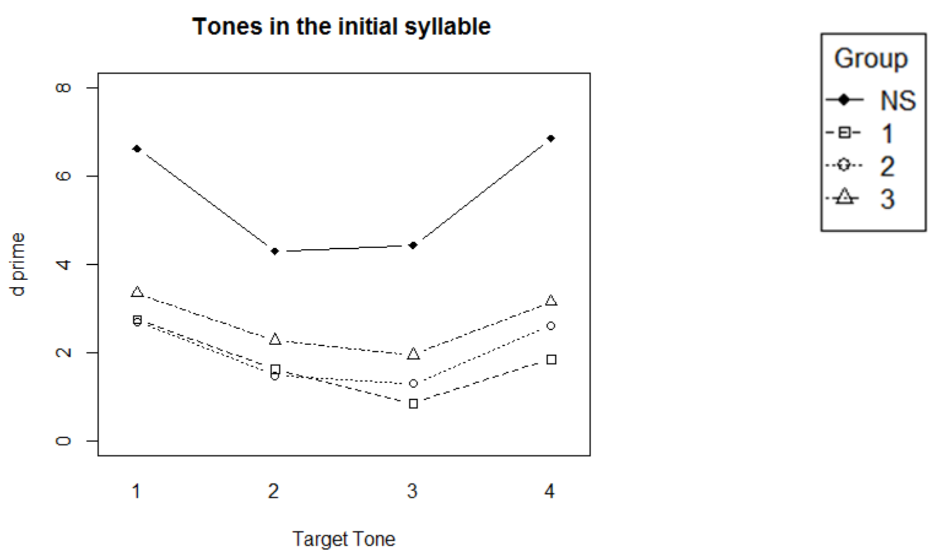


Figure 12: Identification sensitivity of tones in the initial syllable

4.5.2.3 Error Pattern

Table 18 shows the error patterns of Mandarin four tones in the initial syllables with the percentages in parenthesese. Generally, among the learner groups, Tone 1(H) was mostly identified as Tone 2 (LH), Tone 2 (LH) was mostly identified as Tone 3 (L), Tone 3 (L) was mostly identified as Tone 4 (HL), and Tone 4 (HL) was mostly identified as Tone 1 (H). Compared with the learner groups, Group NS had different error patterns. Specifically, Tone 3 (L) was mostly identified as Tone 2 (LH), and Tone 4 (HL) was mostly identified as either Tone 1(H) or Tone 3 (L).

A two-way ANOVA results showed that the effect of group was significant ($F(3,33)= 4.350$, $p<0.05$), and the error of learner groups were significnatly higher than Group NS. Generally, among the learner groups, with more learning experience there were less errors in tone identification. The effect of error pattern was significant ($F(11,33)=5.947$, $p<0.001$). Error Tone 2 (LH)-to-Tone 3 (L) (which indicates the target syllable Tone 2 (LH) being identified as Tone 3 (L)) ($t(33) = 2.711$, $p<0.05$), Tone 3 (L)-to-Tone 2 (LH) ($t(33) = 2.608$, $p<0.05$) and Tone 3 (L)-to-Tone 4 (HL) ($t(33) = 3.358$, $p<0.01$) were significantly higher than other error patterns.

Table 18: Confusion matrix for tones in the initial syllable with the percentages in parentheses

Group	Stimulus	Response				
		Tone 1 (H)	Tone 2 (LH)	Tone 3 (L)	Tone 4 (HL)	NA
Group 1	Tone 1 (H)	**121 (68.8)**	17 (9.7)	4 (2.3)	18 (10.2)	16 (9.1)
	Tone 2 (LH)	13 (7.4)	**93 (52.8)**	46 (26.1)	11 (6.3)	13 (7.4)
	Tone 3 (L)	3 (2.2)	25 (18.9)	**42 (31.8)**	48 (36.4)	14 (10.6)
	Tone 4 (HL)	17 (9.7)	14 (8.0)	6 (3.4)	**128 (72.7)**	11 (6.25)
Group 2	Tone 1 (H)	**122 (69.3)**	22 (12.5)	9 (5.1)	17 (9.7)	6 (3.4)
	Tone 2 (LH)	26 (14.8)	**94 (53.4)**	34 (19.3)	5 (2.8)	17 (9.7)

	Tone 3 (L)	4 (3.03)	31 (23.5)	**61 (46.2)**	24 (18.2)	12 (9.1)
	Tone 4 (HL)	12 (6.8)	14 (7.9)	9 (5.1)	**131 (74.4)**	10 (5.7)
Group 3	Tone 1 (H)	**157 (81.8)**	15 (7.8)	1 (0.5)	13 (6.8)	6 (3.1)
	Tone 2 (LH)	17 (8.9)	**114 (59.4)**	34 (17.7)	9 (4.7)	18 (9.4)
	Tone 3 (L)	4 (2.7)	30 (20.8)	**62 (43.1)**	37 (25.7)	11 (7.6)
	Tone 4 (HL)	18 (9.4)	3 (1.6)	6 (3.1)	**158 (82.3)**	7 (3.6)
Group NS	Tone 1 (H)	**175 (91.1)**	7 (3.6)	4 (2.1)	4 (2.1)	2 (1.0)
	Tone 2 (LH)	9 (4.7)	**157 (81.8)**	19 (9.9)	4 (2.1)	3 (1.6)
	Tone 3 (L)	2 (1.4)	12 (8.3)	**126 (87.5)**	3 (2.1)	1 (0.7)
	Tone 4 (HL)	6 (3.1)	3 (1.6)	6 (3.1)	**176 (91.7)**	1 (0.5)

Concerning error patterns of Tone 3 in the initial syllable in Tone 3-Tone 3 combination, see Table 19. It was found that there was no group difference between learner groups and Group NS ($F(3, 9) = 0.381$, $p = 0.769$). Both learner groups and Group NS identified Tone 3 in Tone 3-Tone 3 combination as Tone 2 over 50%. This confirmed that the derived Tone 2 (LH) from Tone 3 (L) and the underlying Tone 2 (LH) were perceptually indistinguishable to Mandarin speakers (Wang & Li, 1967; Peng, 1996).

Table19: Confusion matrix for Tone 3 in the initial syllable in Tone 3-Tone 3 combination with the percentages in parentheses.

Group	Stimulus	Response				
		Tone 1 (H)	Tone 2 (LH)	Tone 3 (L)	Tone 4 (HL)	NA
Group 1	Tone 3	4 (9.09)	24 (54.5)	**8 (18.2)**	1 (2.3)	7 (15.9)
Group 2	Tone 3	2 (4.5)	23 (52.3)	**11 (25.0)**	4 (9.1)	4 (9.1)

Group 3	Tone 3	5 (10.4)	25 (52.1)	**11 (22.9)**	3 (6.3)	4 (8.3)
Group NS	Tone 3	2 (4.2)	31 (64.6)	**13 (27.1)**	1 (2.1)	1 (2.1)

4.5.2.4 Reaction Time

According to the descriptive statistics, the reaction time for identification of the four Mandarin lexical tones in the initial syllables, Group NS had the least reaction time compared with the other three learner groups. The reaction time for different groups from short to long in terms of miliseconds was Group NS (899.435) < Group 3 (1218.377) < Group 1 (1249.850) < Group 2 (1290.536). For Group NS, the reaction time of the four Mandarin tones from short to long was Tone 4 (818.157) < Tone 1 (906.909) < Tone 2 (926.796) < Tone 3 (968.495). For learner groups, the reaction time from short to long was Tone 4 (1176.581) < Tone 1 (1178.274) < Tone 3 (1367.738) < Tone 2 (1385.659). Generally, Tone 1 (H) and Tone 4 (HL) took shorter time to identify than Tone 2 (LH) and Tone 3 (L) for both Group NS and learner groups, as shown in Figure 13.

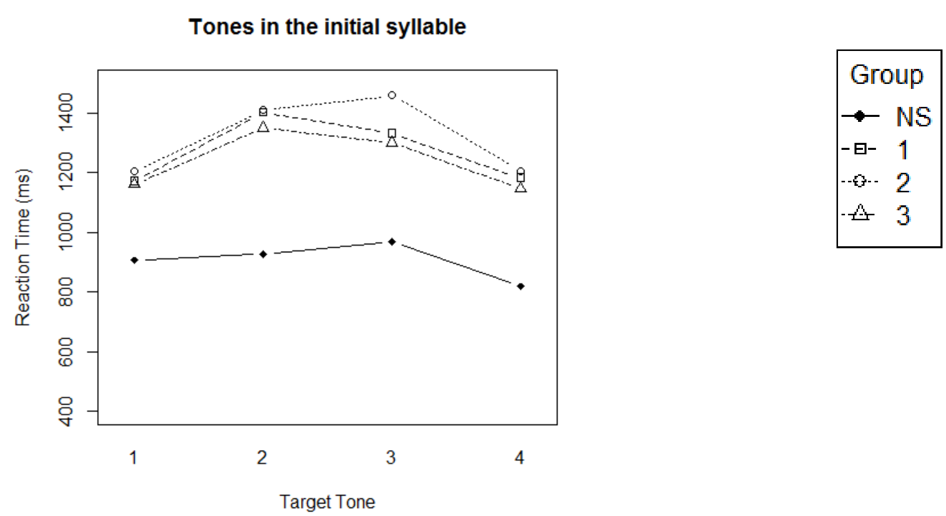


Figure 13: Reaction time of tones in the initial syllable

LMM was used to regress the identification reaction time for the tones in the initial syllable of each trial for group, tone, and the interactions between them. The full model showed that there was significant main effect of group ($F(3, 42) = 10.325$, $p < 0.001$) and tone ($F(3, 1865) = 30.190$, $p < 0.001$), as well as significant interaction between group and tone effects ($F(9, 1865) = 1.943$, $p < 0.05$). Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that the L2 learner groups' reaction time in identification was not significantly differently among each other, while they were all significantly different from Group NS ($p < 0.001$).

We futher explored the context effect including compatible and conflicting contexts on idenficiation reaction time of tones in the initial syllable (see Figure 14). According to the descriptive statistics, the reaction time was shorter in compatible context than in conflicting context, including learner groups on average: compatible context (1220.640) < conflicting context (1281.559) and Group NS: compatible context (894.348) < conflicting context (903.947). LMM showed that there was significant context effect ($F(1, 1869) = 8.959$, $p < 0.01$) and group effect ($F(3, 40) = 9.261$, $p < 0.001$). The group and context interaction effect was significant ($F(1, 1869) = 7.619$, $p < 0.01$). The reaction time of all the learner groups were significantly longer than Group NS. The tone identification reaction time in the conflicting context was significantly longer than its in the compatible context.

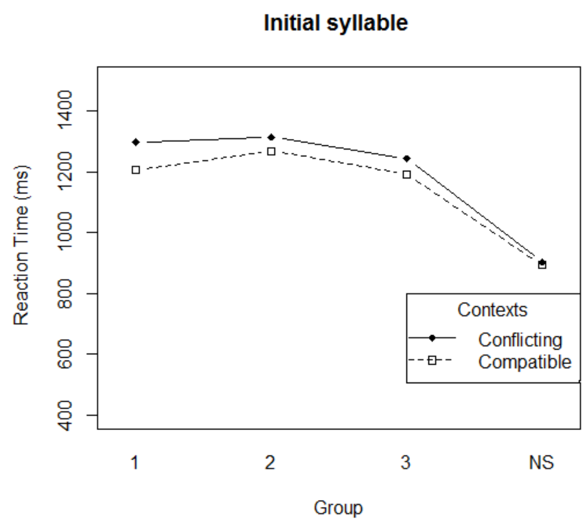


Figure 14: Context effects on reaction time in the initial syllable

To summarize, learner groups' accuracy and identification sensitivity was significantly lower and their identification reaction time was significantly slower than native Mandarin speakers. L2 Learners with more learning experience were more resemblant to those of native speakers of Mandarin, though L2 learner groups did not perform statistically differently among each other. For learner groupsn, Tone 3 (L) was with the lowest acuracy and lowest identification sensitivity. In contrast, Tone 2 (LH) was with the lowest acuracy and lowest identification sensitivity for Group NS. In conflicting context, the accurracy was significantly lower and the identification reaction time was significantly slower than its in compatible context. No statistical significance was found for Group NS. Tone 1 (H) and Tone 4 (HL) took shorter reaction time to identify than Tone 2 (LH) and Tone 3 (L) for both Group NS and learner groups. Among learner groups, Tone 1 (H) was mostly identified as Tone 2 (LH), Tone 2 (LH) was mostly identified as Tone 3 (L), Tone 3 (L) was mostly identified as Tone 4 (HL), and Tone 4 (HL) was mostly identified as Tone 1 (H). Error Tone 2 (LH)-to-Tone 3 (L) (which indicates the target syllable Tone 2 (LH) being identified as Tone 3 (L)), Tone 3 (L)-to-Tone 2 (LH) and Tone 3 (L)-to-Tone 4 (HL) were significantly higher. In addition, there was no significant group difference concerning tone identification of Tone 3 in Tone 3-Tone 3 combination between learner groups and Group NS.

4.5.3 Experiment 3: The Final Syllable of Disyllabic Non-words

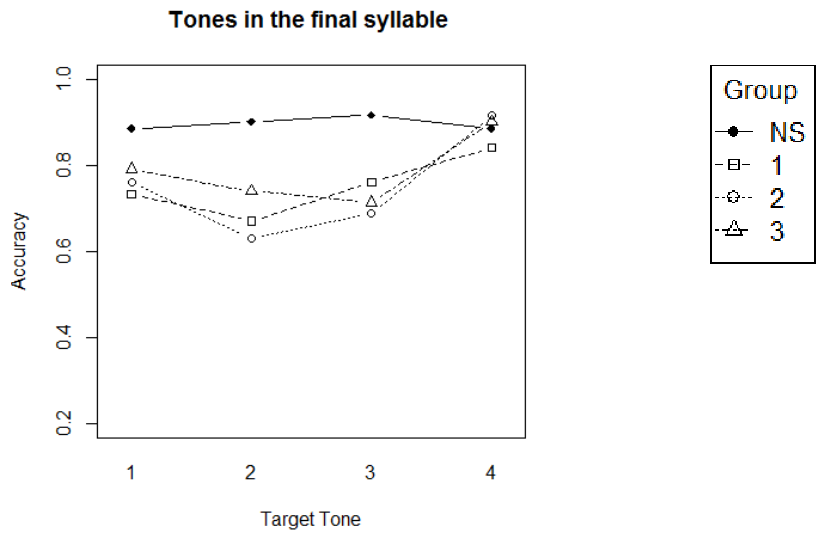
4.5.3.1 Accuracy Rate

Figure 15 shows the four groups' accuracy of the four Mandarin lexical tones in the final syllable of disyllabic non-words.

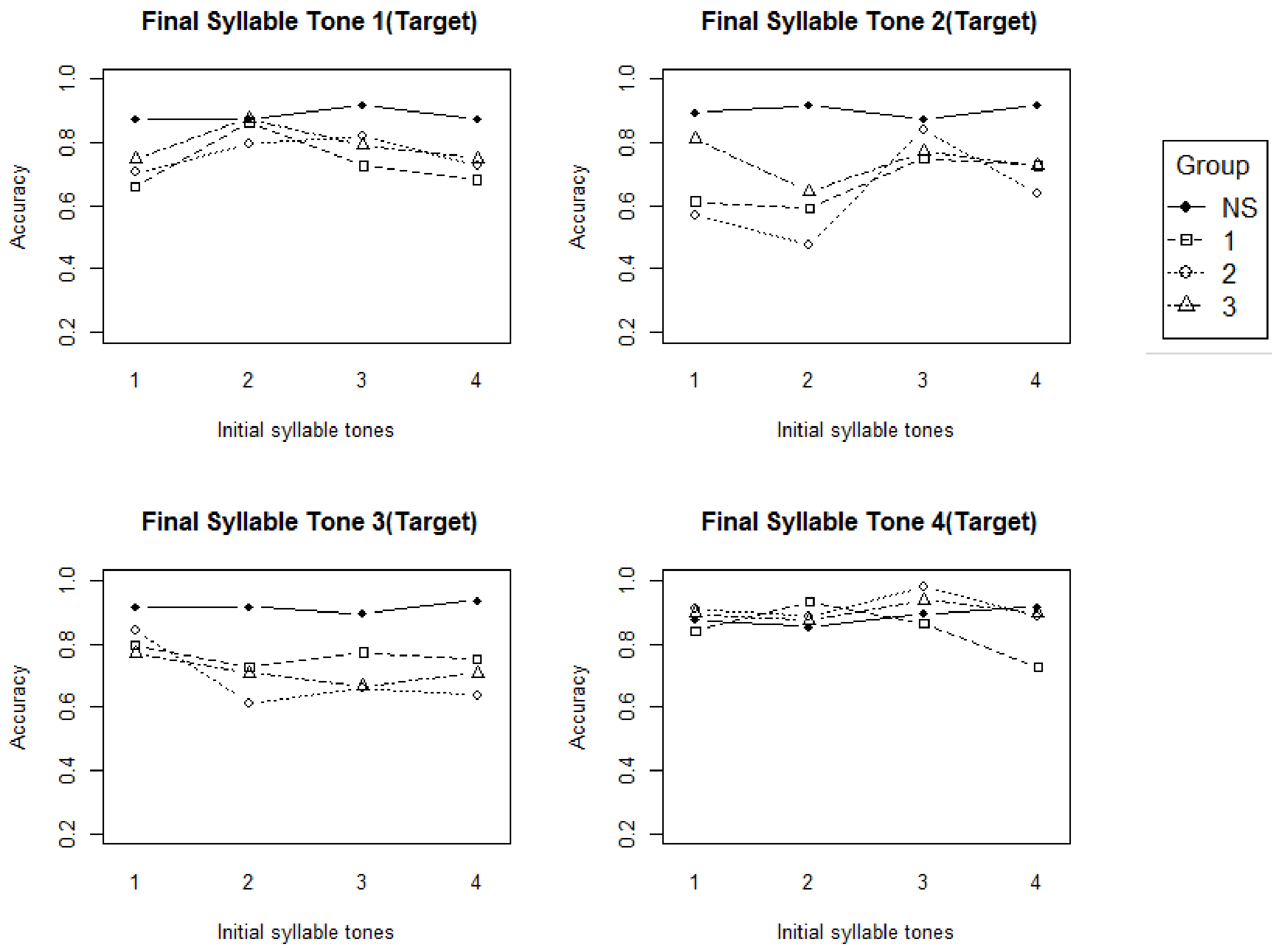
According to the descriptive statistics, the accuracy mean of Group NS was the highest compared with the other three learner groups. The accuracy mean of different groups from high to low was Group NS (0.897) > Group 3 (0.786) > Group 1 (0.751) > Group 2 (0.748). For learner groups, the accuracy mean of tones from high to low was Tone 4 (0.886) > Tone 1 (0.762) > Tone 3 (0.721) > Tone 2 (0.682). For Group NS, the accuracy mean of tones from high to low was Tone 3 (0.917) > Tone 2 (0.901) > Tone 1 (0.885) = Tone 4 (0.885).

The GLMM was used to regress the accuracy of tones in the final syllable of each trial (1 for correct and 0 for incorrect) for group, target tones in the final syllable, the tones in the initial syllable and the interactions between target tones and group as well as between target tones and the tones in the initial syllable. The models showed that there was significant main effect of group (Wald $\chi2$ (3)) = 16.924, $p < 0.001$) and significant interaction effect between target tones and group (Wald $\chi2$ (9)) = 33.164, $p < 0.001$). Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that the L2 learner group did not perform statistically differently among each other, while they were all significantly different from Group NS ($p <0.05$). Specifically, it was found that when the final syllable Tone 1 (H), Tone 2 (LH) and Tone 3 (L) was the target, learner groups were all found to be significantly different from Group NS. However, no significant difference between learner groups and Group

NS was found when the final syllable Tone 4 (HL) was the target tone. The accuracy of Tone 2 (LH) ($z = -2.916$, $p<0.01$) was significantly lower and Tone 4 (L) ($z = 5.158$, $p<0.001$) was significantly higher, according to the estimated regression coefficients along with Z statistics.



(a)



(b)

Figure 15: Accuracy rate of the tones in the final syllable

To further explore how tone identification accuracy in the final syllable was influenced by both tones in the initial syllable and final syllable, we futher explored the context effect including compatible and conflicting context on the tone identification accuracy in the final syllable. According to the descriptive statistics, the identification accuracy was higher in compatible context than in conflicting context only for learner groups: compatible context (0.788) > conflicting context (0.743). For Group NS : conflicting context (0.905) > compatible context (0.886). GLMM showed that there was significant group effect (Wald $\chi 2$ (3) = 15.020, $p < 0.01$) and context effect was marginally significant (Wald $\chi 2$ (1) = 3.315, $p = 0.069$). In addition, it was found that with additional learning experience, the effect of the compatible/conflicting context decreases (see Figure 16), as the accuracy mean differences between compatible and conflicing context learner groups declined.

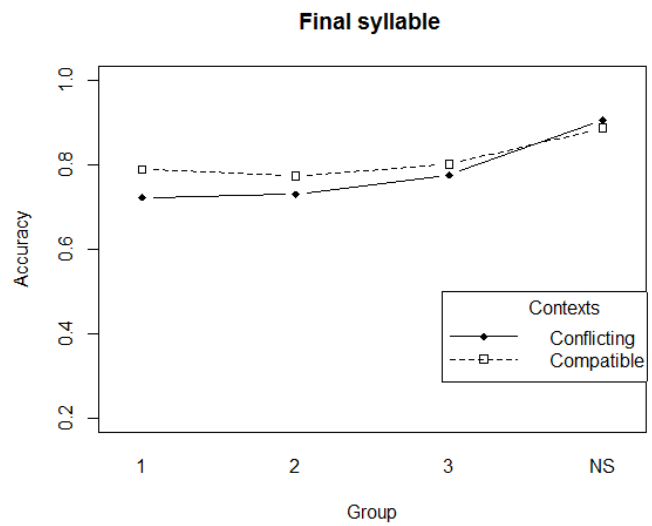


Figure 16: Context effects on accuracy in the final syllable

4.5.3.2 Identification Sensitivity (d prime)

Figure 17 shows the identification sensitivity of the four Mandarin lexical tones in the final syllable.

According to the descriptive statistics, Group NS had the highest identification sensitivity compared with the other three learner groups. The identification sensitivity of different groups from high to low was Group NS (5.696) > Group 3 (3.504) > Group 1 (3.156) > Group 2 (3.007). For learner groups, the identification sensitivity regarding tones in the final syllable from high to low was Tone 4 (4.161) > Tone 1 (3.316) > Tone 3 (2.864) > Tone 2 (2.581), which is similar to the result in the initial syllable. For Group NS, the identification sensitivity of tones in the final syllable from high to low was Tone 3 (6.480) > Tone 4 (5.813) > Tone 1 (5.502) > Tone 2 (4.989). When comparing Figure 17 with Figure 15 (a), namely calculating identification sensitivity versus solely calculating identification accuracy, the identification sensitivity was more closely associated with learning experience which indicates that identification sensitivity has better discriminating power than identification accuracy with less overlap among the learner groups compared with the identification accuracy results.

LMM was used to regress the identification sensitivity for the tones in the final syllable of each trial for group, tone, and the interactions between them. The full model showed that there was neither significant main effect nor significant interaction effect. Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that the L2 learner groups' identification sensitivity did not significantly differ among each other, while they were all significantly different from Group NS ($p < 0.05$).

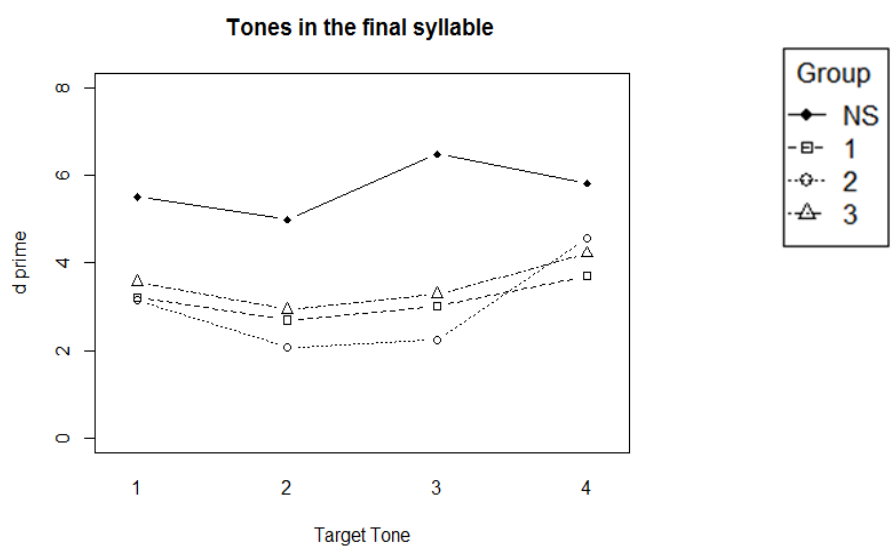


Figure 17: Identification sensitivity of tones in the final syllable

4.5.3.3 Error Pattern

Table 20 shows the error pattern of Mandarin four tones in the final syllable with the percentages in parentheses. Generally among learner groups, Tone 1 (H) was mostly identified as Tone 2 (LH), Tone 2 (LH) was mostly identified as Tone 3 (L), Tone 3 (L) was mostly identified as Tone 2 (LH), and Tone 4 (HL) was mostly identified as Tone 1 (H). For Group NS, Tone 1 (H) was mostly identified as Tone 2 (LH), and Tone 2 (LH) was mostly identified as Tone 3 (L), Tone 3 (L) was mostly identified as Tone 4 (HL), and Tone 4 (HL) was mostly identified as Tone 2.

A two-way ANOVA results showed that the main effect of group ($F(3,33)= 2.994$, $p<0.05$) and error pattern ($F(11,33)= 5.153$, $p<0.001$) was significant. Learner groups were all significant with more errors than Group NS. Generally, with more learning experience there were less errors in tone identification among the learner groups. Error Tone 1 (H)-to-Tone 3 (L) (which indicates the target syllable Tone 1 (H) being identified as Tone 3 (L)), Tone 1 (H)-to-Tone 4 (HL), Tone 2 (LH)-to-Tone 4 (HL), Tone 3 (L)-to-Tone 1 (H), Tone 4 (HL)-to-Tone 1 (H), Tone 4 (HL)-to-Tone 2 (LH), Tone 4 (HL)-to-Tone 3 (L) were significantly lower than other error patterns. In contrast, error Tone 1 (H)-to-Tone 2 (LH), Tone 2 (LH)-to-Tone 1 (H), Tone 2 (LH)-to-Tone 3 (L), Tone 3 (L)-to-Tone 2 (LH), and Tone 3 (L)-to-Tone 4 (HL) were significantly higher.

Table 20: Confusion matrix for tones in the final syllable with the percentages in parentheses

Group	Stimulus	Response				
		Tone 1 (H)	Tone 2 (LH)	Tone 3 (L)	Tone 4 (HL)	NA
Group 1	Tone 1 (H)	**129 (73.3)**	22 (12.5)	6 (3.4)	8 (4.5)	11 (6.3)
	Tone 2 (LH)	18 (10.2)	**118 (67.0)**	23 (13.1)	8 (4.5)	9 (5.1)
	Tone 3 (L)	3 (1.7)	11 (6.3)	**134 (76.1)**	19 (10.8)	9 (5.1)
	Tone 4 (HL)	6 (3.4)	7 (3.9)	4 (2.3)	**148 (84.1)**	11 (6.3)
Group 2	Tone 1 (H)	**134 (76.1)**	20 (11.3)	5 (2.8)	6 (3.4)	11 (6.3)
	Tone 2 (LH)	16 (9.1)	**111 (63.1)**	36 (20.5)	2 (1.1)	11 (6.3)
	Tone 3 (L)	2 (1.1)	37 (21.0)	**121 (68.8)**	11 (6.3)	5 (2.8)

	Tone 4 (HL)	5 (2.8)	3 (1.7)	0	**161 (91.4)**	7 (4.0)
Group 3	Tone 1 (H)	**152 (79.1)**	19 (9.9)	2 (1.0)	8 (4.2)	11 (5.7)
	Tone 2 (LH)	19 (9.9)	**142 (74.0)**	23 (12.0)	3 (1.6)	5 (2.6)
	Tone 3 (L)	2 (1.0)	22 (11.5)	**137 (71.4)**	25 (13.0)	6 (3.1)
	Tone 4 (HL)	8 (4.2)	3 (1.6)	4 (2.1)	**173 (90.1)**	4 (2.1)
Group NS	Tone 1 (H)	**170 (88.5)**	12 (6.3)	2 (1.0)	6 (3.1)	2 (1.0)
	Tone 2 (LH)	3 (1.6)	**173 (90.1)**	8 (4.2)	5 (2.6)	3 (1.6)
	Tone 3 (L)	5 (2.6)	3 (1.6)	**176 (91.7)**	6 (3.1)	2 (1.0)
	Tone 4 (HL)	7 (3.6)	8 (4.2)	4 (2.1)	**170 (88.5)**	3 (1.6)

4.5.3.4 Reaction Time

According to the descriptive statistics, in the final syllables, Group NS had the least reaction time compared with the other three learner groups. The reaction time of different groups from short to long in terms of miliseconds was Group NS (602.517) < Group 3 (760.325) < Group 2 (788.295) < Group 1 (840.167). For Group NS, the tone identification reaction time from short to long was Tone 3 (549.110) < Tone 2 (568.329) < Tone 4 (608.014) < Tone 1 (687.103). For learner groups, Tone 1 (725.478) < Tone 4 (776.907) < Tone 3 (785.118) < Tone 2 (905.142). Generally, Tone 1 (H) and Tone 4 (HL) took shorter reaction time to identify than Tone 2 (LH) and Tone 3 (L) for learner groups but not for Group NS, as shown in Figure 18.

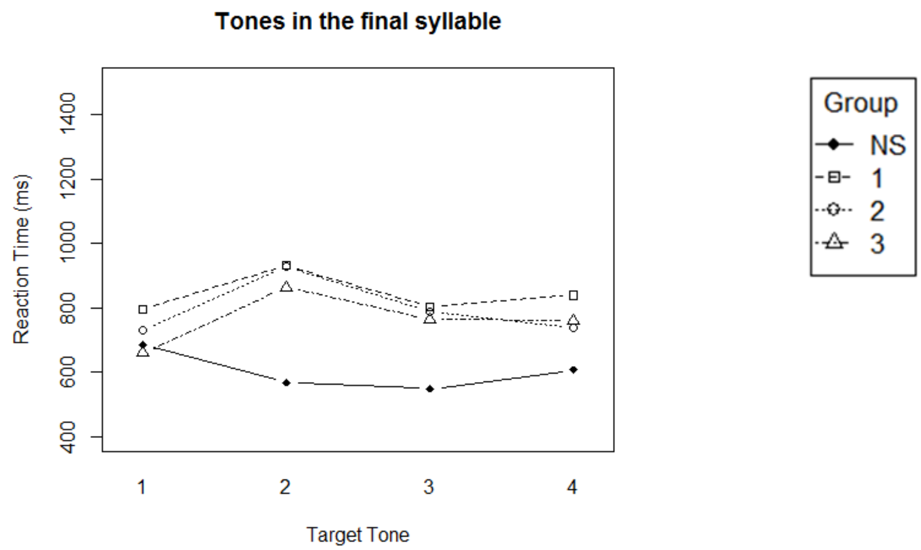


Figure 18: Reaction time of tones in the final syllable

LMM was used to regress the identification reaction time for tones in the final syllable of each trial for group, tone, and the interactions between them. The full model showed that there was neither significant main effect nor interactions. Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that the L2 learner groups' identification reaction time was not significantly differently among each other, while Group 1 ($z=3.614$, $p<0.01$) and Group 2 ($z=2.900$, $p<0.05$) were both significantly slower than Group NS, and Group 3 ($z=2.512$, $p=0.057$) was marginally slower than Group NS.

We futher explored the context effect including compatible and conflicting context on identification reaction time of tones in the final syllable (see Figure 19). According to the descriptive statistics, the reaction time was shorter in conflicting context than in compatible context for Group NS: conflicting context (574.429) < compatible context (639.369) and nearly the same for learner groups on average: compatible context (794.222) and conflicting context (794.999). LMM showed that there was significant group effect ($F (3, 42) = 4.662$, $p < 0.01$) and the reaction time of all the learner groups was significantly longer than Group NS. The interaction between group and context effect was marginally singificant (F (3, 2301) = 2.098, p = 0.098). The main effect of context was not significant (F (1, 2301) = 1.168, p = 0.27).

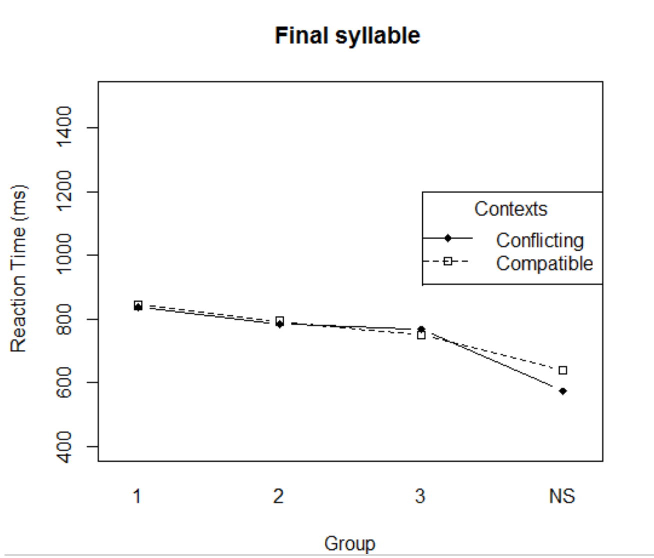


Figure 19: Context effects on reaction time in the final syllable

To summarize, learner groups' accuracy and identification sensitivity was significantly lower and their identification reaction time was significantly slower than native Mandarin speakers. L2 Learners with more learning experience were more resemblant to those of native speakers of Mandarin, though L2 learner groups did not perform statistically differently among each other. Tone 2 (LH) was with the lowest acuracy and lowest identification sensitivity for both learner groups and Group NS. The accuracy of Tone 2 (LH) was significantly lower and Tone 4 (L) was significantly higher. The context effect was not as strong as it was in the initial syllable. The accuracy of learner groups in the conflicting context was marginally lower than in the compatible context. No statistical significance was found for Group NS. No significant difference of context effect was found for identification reaction time. Tone 2 (LH) and Tone 3 (L) took longer reaction time to identify than Tone 1 (H) and Tone 4 (HL) for learner groups but not for Group NS. Among learner groups, Tone 1 (H) was mostly identified as Tone 2 (LH), Tone 2 (LH) was mostly identified as Tone 3 (L), Tone 3 (L) was mostly identified as Tone 2 (LH), and Tone 4 (HL) was mostly identified as Tone 1 (H). Error Tone 1 (H)-to-Tone 2 (LH) (which indicates the target syllable Tone 1 (H) being identified as Tone 2 (LH)), Tone 2 (LH)-to-Tone 1 (L), Tone 2 (LH)-to-Tone 3 (L), Tone 3 (L)-to-Tone 2 (LH), and Tone 3 (L)-to-Tone 4 (HL) were significantly higher.

4.6 Comparison between Tasks

In this section, due to syllable position difference, comparison was made between monosylables, initial syllables of disyllabic nonwords and final syllables of disyllabic non words.

Concerning accuracy, for learner groups, the tone identification accuracy from high to low was monosylable (0.850) > final syllable (0.762) > initial syllable (0.629). For Group NS, monosyllable (0.990) > final syllable (0.897) > initial syllable (0.881). GLMM was used to regress the accuracy for tones in the three syllable positions of each trial (1 for correct and 0 for incorrect) for group, target tones, syllable positions and the interactions between target tones and syllable position. Besides the significant main effects of group (Wald $\chi 2$ (3)) = 28.268, $p < 0.001$) and target tones (Wald $\chi 2$ (3)) = 46.688, $p < 0.001$), there was significant main effect of syllable position (Wald $\chi 2$ (2)) = 65.046, $p < 0.001$) and significant interaction effect between target tone and syllable position (Wald $\chi 2$ (6)) = 72.208, $p < 0.001$). Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that both tone identification accuracy of initial syllable ($z = -15.708$, $p<0.001$) and final syllable ($z = -8.139$, $p<0.001$) was significantly lower than monosyllable. Tone identification accuracy in final syllable was significantly higher than its in initial syllable ($z = 10.375$, $p<0.001$).

Concerning identification sensitivity, d prime, for learner groups, the average tone identification sensitivity from high to low was monosylable (4.535) > final syllable (3.231) > initial syllable (2.173). For Group NS, monosyllable (7.296) > final syllable (5.547) > initial syllable (3.231). LMM was used to regress the identification sensitivity of the three task conditions of each trial for group, tone, syllable position and the interactions between tone and syllable position. The full model showed that there was significant main effect of group (Wald $\chi 2$ (3)) = 38.340, $p < 0.001$), tone (Wald $\chi 2$ (3)) = 90.924, $p < 0.001$), syllable position (Wald $\chi 2$ (2)) = 95.066, $p < 0.001$), and significant interaction effect between tone and syllable position (Wald $\chi 2$ (6)) = 29.207, $p < 0.001$). Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that both tone identification sensitivity of initial syllable ($z = -15.003$, $p<0.001$) and final syllable ($z = -9.430$, $p<0.001$) was significantly lower than monosyllable. Tone identification sensitivity in the final syllable was significantly higher than it in the initial syllable ($z = 5.603$, $p<0.001$).

Concerning error patterns, a two-way ANOVA tested among the learners to see how the syllable position influence the error patterns. The results showed that the main effect of group ($F(3,105)= 12.465$, $p<0.001$), error pattern ($F(11,105)= 13.849$, $p<0.001$), and syllable position ($F(2,105)= 27.021$, $p<0.001$) was significant. Also, there was significant interaction effect between error pattern and syllable position ($F(22,105)= 2.587$, $p<0.001$). Pairwise comparisons of estimated regression coefficients along with Z statistics, by using Tukey HSD, showed that both error rate of the initial syllable ($z = 6.463$, $p<0.001$) and final syllable ($z = 2.649$, $p<0.05$) was significantly higher than its in monosyllables. Error rate in the final syllable were significantly lower than its in the initial syllable ($z = -3.913$, $p<0.001$).

Concerning identification reaction time of tones in the three syllable conditions, for learner groups, monosyllable (588.136) < final syllable (738.293) < initial syllable (1134.482). For Group NS, monosyllable (473.584) < final syllable (602.516) < initial syllable (899.435). No significant main effect or interaction effect was found concerning tone, group and syllable position.

4.7 Discussion

In the current study, it was investigated how tonal context and syllable position would affect L1 English learners of Chinese' tone identification accuracy, sensitivity, error patterns, and reaction time and whether their identification of Mandarin tones in monosyllables and disyllabic non words improve with more Mandarin learning experience. The acoustic analysis of tone production in context, successfully predicts the difficulty in identifying Mandarin tones in context.

The major findings of this study are as follows:

As expected, both syllable position and context affects the tone identification of L1 English learners of Chinese. Tones in monosyllables were identified with the highest accuracy and sensitivity, and shortest reaction time, followed by tones in the final syllable and tones in the initial syllable. It was found that syllable position has a significant effect on accuracy, sensitivity, and error pattern. Additionally, fewer errors were made in the compatible context than the conflicting context, which supports Xu (1993, 1994). With more learning experience, the effect of the compatible/conflicting context decreased for both tones in the initial syllable and final syllable tasks. However, no statistical significance was found for Group NS.

The identification accuracy and sensitivity of Tone 1 (H) and Tone 4 (HL) was better than Tone 2 (LH) and Tone 3 (L) among the three experimental conditions, including monosyllables, the initial syllable and the final syllable of disyllabic non-words, which reconfirmed previous studies (Whalen & Xu, 1992; Wang & Li, 1967; Peng, 1996). It was found that Tone 4 was the most easily identified tone when presented in both monosyllables and in the final syllable, which supports a previous study Broselow et al. (1987). Also, the present study supports that Tone 2 (LH) is possibly the most challenging tonal category for L1 English speakers to perceive (Lee et al., 2010), as it is consistently confused with Tone 3 (L) (Gottfried & Suiter, 1997; Kiriloff, 1969; Miracle, 1989; Shen, 1989; Hao, 2012). As expected, the confusion between Tone 2 (LH) and Tone 3 (L) was most salient, among the three experimental conditions. In terms of error patterns, it was as expected that in the initial syllable, error Tone 2 (LH)-to-Tone 3 (L) (which indicates the target syllable Tone 2 (LH) being identified as Tone 3 (L)), Tone 3 (L)-to-Tone 2 (LH) and Tone 3 (L)-to-Tone 4 (HL) was significantly higher. In the final syllable, it was as expected that error Tone 1 (H)-to-Tone 2 (LH), Tone 2 (LH)-to-Tone 1 (H), Tone 2 (LH)-to-Tone 3 (L), Tone 3 (L)-to-Tone 2 (LH) was significant higher. Without being expected, Tone 3 (L)-to-Tone 4 (HL) was also significantly higher in the final syllable. Error between Tone 2 (LH) and Tone 4 (HL), though expected, was not found in the present study.

Learner groups were significantly different from Group NS, in terms of tone identification accuracy, sensitivity, error patterns, and reaction time. L2 Learners with more learning experience were more resemblant to those of native speakers of Mandarin, though L2 learner groups did not perform statistically differently among each other. In addition, the reaction time of Tone 3 (L) for Group NS was shortest in both monosyllable and tones of the final syllable, indicating that Group NS may have employed creaky voice quality to identify Tone 3 (L) (Gottfried & Suiter, 1997).

Several implications of the findings are further discussed below:

First, it was observed that tone identification of the final syllable was better and tone identification of the initial syllable was more challenging for L2 learners, which is similair to the results found in Gottfried & Suiter (1997) and He (2010). One of the explanations for the finding is that L2 learners give their primary perceptual cue to the F0 offset value of the tone (Gilber & Liu, 2013; Liu, 2013), and the magnitude of the variation might be less relavent. This is supported by relating the tone identification result from Study II to the tone production result in Study I. In Study I, the F0 value of of the tone onset in the final syllables deviated more from the canonical form than the F0 value of the tone offset in initial syllable, however the offset F0 value of the tone in the final syllable was intact, which may potentially makes it easier for listeners to identify the tones in the final syllable. In addition, F0 values in the initial syllable were less correlated to the related F0 value in monosyllables and the offset F0 value in the initial syllable was influenced by the F0 value in the final tone, which brought more challenge to the tone identification tasks of the initial syllable for L2 learners. Another explanation is the nature of the initial and final syllable position in perception studies. Lin & Wang (1985) found that tones are perceived relative to other tones and the offset of the target tone assimilates to the onset of the following syllable tone perceptually, which may possibly make it harder to identify the tone in the initial syllable. When the following syllable tone is the perception target, it is found that the initial tone functions as a reference tone and the effect is contrastive (Gårding et al., 1986; Huang & Holt, 2008), which may possibly make it easier to identify the tone in the final syllable. Also, it could be possible that final syllable has the advantage in acoustic memory. According to Darwin & Baddeley (1974), acoustically distinct consonants in syllable-final position received large and highly significant recency and modality effects. This gives evidence to acoustic memory contributing to recollection (Darwin & Baddeley, 1974), therefore it is possible that the tone in the final syllable is lastly heard and therefore has the advantage over the tones in the initial syllable.

Second, no statistical significance of the context effect was found for Group NS, which reconfirmed that native Mandarin-speaking listeners perceive Mandarin lexical tones categorically (Stagray & Downs, 1993; Hallé et al., 2004; Xu et al., 2006; Xi et al., 2010), as they exhibited the decreased within-category sensitivity. On the contrary, among learner groups, identification accuracy of tones in the compatible context was higher than in the conflicting context, though the context effect on tone identification in the final syllable was not as strong as it was in the initial syllable, and the phonetic variation caused by tonal contexts presents a substantial difficulty for learner groups. This supports that nonnative listeners, who have not acquired the phonetic boundaries between tones, may misperceive any within categorical phonetic variation as linguistically relevant (Stagray & Downs, 1993; Mattock, 2006; Huang & Johnson, 2010; Chang, 2011). Also, among the L2 learner groups, with more learning experience, the effect of the compatible/conflicting context decreased for both tones in the initial syllable and final syllable tasks, indicating that learning experience reduces the linguistically irrelevant acoustic input. The result found in the present study supports NLM (Kuhl & Iverson, 1995), which suggests that exposure to language early in life produces a change in perceived distances in the acoustic space underlying phonetic distinctions. With more learning experience, perceptual distance between the prototype and its surrounding stumili is shrunk, while the region near the phonetic boundary is perceptually streched (Kuhl & Iverson, 1995), therefore the effect of the compatible/conflicting context decreased for the identification of tones in context among L2 learner groups.

Third, the present study supports that Tone 2 (LH) is possibly the most challenging tonal category for L1 English speakers to perceive (Lee et al., 2010), as it is consistently confused with Tone 3 (L) (Gottfried & Suiter, 1997; Kiriloff, 1969; Miracle, 1989; Shen, 1989; Hao, 2012). The identification sensitivity of Tone 2 (LH), which gives a more comprehensive perspective, was the lowest in both monosyllables and the final syllable among L2 learner groups. It is highly possible that the phonetic variations of Tone 2 (LH) in context can be resemblant to both Tone 1 and Tone 3, in terms of F0 height and F0 contour, which makes it the hardest tone to identify. Tone 3 is possibly the most challeging tonal category for L2 learners to identify in a non final syllable position. The confusion between Tone 2 (LH) and Tone 3 (L) was most salient, among the three experimental conditions. It may be due to the complicated phonological relationship between Tone 2 (LH) and Tone 3 (Hao, 2012). For the phonological tone sandhi of Tone 3, the derived Tone 2 (LH) from Tone 3 (L) and the underlying Tone 2 (LH) were perceptually indistinguishable even to Mandarin speakers (Wang & Li, 1967; Peng, 1996). Drawing comparison to the identification accuracy of native speakers, for Group NS, the accuracy was also very low when the tone combination was Tone 2 (LH)-Tone 3 (L) or Tone 3 (L)-Tone 2 (LH), which indicates the identification difficulty of Tone 2 and Tone 3 in context in nature, due to the acoustic similarities (Moore and Jongman, 1997; Hao, 2012) and supports that these two tones can sometimes also confuse native listeners (Shen & Lin, 1991). In addition, the identification difficulty of Tone 3 can be due primarily to the difficult nature of perceiving Tone 3 (Moore and Jongman, 1997) and secondly to half third tone sandhi. When the Tone 3 in the initial syllable, all the learner groups mostly identified it as Tone 4, which indicates that a perception realization of half third tone falling exists and this may be due to the greater emphasize on the tonal contour, and lacking in appropriate attention to the starting point. It would be effective to have L2 learners perceive the contrast of Tone 4 and the half third tone to have them capture the intrinsic perceptural cues to from the tonal categories during an intensive training procedure (Wang, et al, 1999, 2003). Error between Tone 2 (LH) and Tone 4 (HL), though expected, was not found in the present study, which indicates that learners with more langauge learning experience later placed emphasis on the tone direction (Gandour, 1978, 1983, 1984).

Fourth, the present study illustrated the substantial difficulty for L1 learners of English in Tones identification, due to tonal context, syllable position and learning experience. The pedagogical implications from this study can benefit the L2 language teachers in predicting the points of difficulty in learning Mandarin tones and in helping L2 learners improve their tone identification accuracy. To improve tone identification, Chinese language teachers should emphasize the tone perception/identification continuously for high level L2 learners. In addition, the identification of Mandarin tones need to be trained in context, as the tones in non-final position and in conflicting context are harder and should be given more attention. The tone pairs of Tone 3 (L/dipping) and Tone 2 (LH), as well as the half Tone 3 (L) and Tone 4 (HL) need to be given more attention. In addition, language teachers should exaggerate the acoutic differences between the confused tone pairs, as it has been observed in Infant-Directed Speech (IDS) that Mandarin-speaking infants' care giver exaggeration of the acoutic differences in IDS was obserbed and IDS does not distort the acoustic cues that are essential to word meaning at the syllable level, which may shed light on the instructional strategy in L2 acuiqisition (Liu et al., 2007). Therefore, after concentrated training regarding problem areas, students will exponentially improve their Chinese tone production and identification.

Concerning limitations, the current study only looked at nonsense words and tone combinations /ma.ya/ and /ya.ma/, the simplest condition, just concerning itself with acoustic input and how it influences L1 English learners of Chinese's identification difficulty, excluding many other influences, including vowels, lexical frequency, male and female speakers, etc. To make the study more comprehensive, other issues should be explored as well. E.g. more vowels should be included to enlarge the number of the stimuli and see how that may affect non-native listeners' Mandarin tone perception/identification, as different vowels will provide varied tonal F0 patterns (Howie, 1976). In addition, the present study only explored L2 learners' identification of Mandarin tones, without their production information, we cannot see the full picture of the second language acquisition of Mandarin tones.

For follow-up studies, a possible study could be how L1 English learners of Chinese' tone identification can be influenced by perception of meaningful Chinese words, and how it is different from perceiving the nonsense words. Also, it can be tested how speaking speed will influence non-native learners' tone identification, by recording stimuli with different speeds, including rapid and normal. It would also be interesting to explore the tone identification of L1 English learners of Chinese who study in the target language environment, i.e. how different learning environments will influence students' tone identification, and how their tone perception might different. Also, it would be worthwhile to examine the tone production of naïve speakers and speakers with first year Chinese background to explore the correlations between their tone perception/ identification and production.

CHAPTER 6

CONCLUSION

In modern language use, most Mandarin Chinese words have disyllabic construction (Duanmu, 1999). Current studies explored native Mandarin speakers' production (Study I) and L2 learners' identification (study II) of Mandarin tones presented in monosyllables and two tone sequence with a focus on the influence of tonal contexts.

In Study I, the native Mandarin speakers' production of two tone sequences of Mandarin Chinese has been systematically examined from perspectives of both compatible and conflicting contexts, as well as carry over and anticipatory effects. It has been found that F0 contour and F0 height of the tone in both initial and final syllables was different from the F0 value of the tone in monosyllables. The F0 value of the final syllables deviated more from the canonical form than their counterparts in the initial syllable and the F0 value in the final syllables was significantly more susceptible to the influence of the tonal context. F0 values in the initial syllables were less related to the related F0 value in monosyllables, and F0 values in the final syllables were more correlated to the related F0 value in monosyllables.

These tonal variations in context brought more challenge to the tone identification tasks for both L2 learners and native Mandarin speakers, which was tested among the three experimental conditions, including monosyllables, the initial syllable and the final syllable of disyllabic non-words in Study II. It was found that the tone identification accuracy rates, identification sensitivity, error patterns, and reaction times were significantly influenced by tone, syllable position, tonal context, and learning experience, among L2 learners. Specifically, tones in monosyllables were identified with the highest accuracy and sensitivity, and shortest reaction time, followed by tones in the final syllable and tones in the initial syllable. Fewer errors were made in the compatible context than the conflicting context. With more learning experience, the effect of the compatible/conflicting context decreased for both tones in the initial syllable and final syllable tasks. L2 Learners with more learning experience were more resemblant to those of native speakers of Mandarin, though L2 learner groups did not perform statistically differently among each other. The identification accuracy and sensitivity of Tone 1 (H) and Tone 4 (HL) was better than Tone 2 (LH) and Tone 3 (L) among the three experimental conditions. The confusion between Tone 2 (LH) and Tone 3 (L) was most salient, among the three experimental conditions.

These findings suggest that L2 learners may give their primary perceptual cue to the F0 offset value of the tone (Gilber & Liu, 2013; Liu, 2013) and the magnitude of the variation might be less relevant. L2 learners, who have not acquired the phonetic boundaries between tones, may misperceive any within categorical phonetic variation as linguistically relevant (Stagray & Downs, 1993; Mattock, 2006; Huang & Johnson, 2010; Chang, 2011). In addition, the study suggests that Chinese language teachers should train L2 leaners' to identify/perceive Mandarin tones in context and give more attention to distinguishing the tone pairs of Tone 3 and Tone 2, as well as the half Tone 3 and Tone 4.

The current research helps fill the gap in knowledge about L2 learners' identification difficulty of two tone sequences of Mandarin lexical tones, caused by the acoustic variation existing in native Mandarin speakers' production. This new information contributes to a deeper understanding of context effects on Mandarin tone identification of L2 learners.

References

Baayen, R. H. (2008). *Analyzing linguistic Data, A practical introduction to using R.* Cambridge University Press.

Bates, D., Maechler, M. , Bolker, B., Walker, S. (2014). Linear mixed-effects models using Eigen and S4. [ftp://131.211.84.186/mirror/CRAN/web/packages/lme4/lme4.pdf]

Bent, T. (2005). *Perception and Production of Non-Native Prosodic Categories.* Dissertation. Northwestern University.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research* (171-204).

Blicher, D. L., Diehl, R. L., and Cohen, L. B. (1990). Effects of syllable duration on the perception of the Mandarin tone2/tone3 distinction: Evidence of auditory enhancement, *Journal of Phonetics***,** 18, 37–49.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International* **5:9/10**, 341-345.

Broselow, E., Hurtig, R. R., & Ringen, C. (1987). The perception of second language prosody. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage Phonology: The Acquisition of a Second Language Sound System* (350–362). Cambridge, MA: Newbury House Publishers.

Byrd, D., Lee, S., Riggs, D., & Adams, J. (2005). Interacting effects of syllable and phrase position on consonant articulation. *Journal of the Acoustical Society of America*, 118, 3860–73.

Cao, W. (2010). Mandarin tone perception：A report on level F0 contours. *Zhong Guo Yu Wen (中国语文)*, 6: 536-543.

Chao, Y. R. (1930). A System of Tone Letters. *Le Maître phonétique*, 45:24-27.

Chao, Y. R. (1968). *A Grammar of Spoken Chinese.* Berkeley& Los Angeles: University of California Press.

Chang, Y. H. (2008). Rethinking tones in the Chinese classroom: A design for a beginner's course in Chinese pronunciation. MA Thesis, Indiana University-Bloomington.

Chang, Y.-H. (2011). Distinction between Mandarin Tones 2 and 3 for L1 and L2 listeners. *Proceedings of the 23rd North American Conference on Chinese Linguistics*, 2011 Volume 1, 84-96.

Chen, G. (1974). Pitch Range of English and Chinese Speakers. *Journal of Chinese Linguistics,* 2(2): 159-171.

Chen, Q. (1997). Toward a sequential approach for tonal error analysis. *Journal of the Chinese Language Teachers Association*, 32 (1), 21-39.

Chen, T. H. & Massaro, D. W. (2008). Seeing pitch: Visual information for lexical tones of Mandarin-Chinese. *Journal of the Acoustical Society of America*, 123, 2356 – 66.

Darwin, D. J., & Baddeley, A. D. (1974). Acoustic Memory and the Perception of Speech. *Cognitive Psychology*, 6, 41-60.

Duanmu, S. (1999). Stress and the Development of Disyllabic Words in Chinese. *Diachronica,* XVI: 1.1-35.

Duanmu, S. (2000). "Tone: An overview," in L. L.-S. Cheng and R. Sybesma (eds.). *The First Glot International State-of-the-Article Book: The Latest in Linguistics*, Mouton de Gruyter, Berlin, 251-286.

Duanmu, S. (2007). *The phonology of standard Chinese (2nd Ed.)*. New York, NY: Oxford University Press.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (233-277). Timonium, MD: York Press.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35,* 116-124.

Fox, R., & Qi, Y. Y. (1990). Context effects in the perception of lexical tone. *J. Chinese Ling*, 18, 261-283.

Fu, Q.-J., & Zeng, F.-G. (2000). Identification of temporal envelope cues in Chinese tone recognition. *Asian Pacific Journal of Speech, Language and Hearing*, 5, 45-57.

Gandour, J. T. (1978). The perception of Tone. *Tone: A linguistic Survey, edited by V. Fromkin.* New York: Academic Press, 41-76.

Gandour, J. T. (1983). Tone perception in far Eastern languages. *Journal of Phonetics*, 11: 149-175.

Gandour, J. T. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics*, 12, 235-261.

Gandour, J. T., & Harshman, R. (1978). Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and Speech,* 21, 1-33.

Gandour, J. T., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., Li, X., & Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *Neuroimage,* 23:344–357.

Gandour, J. T., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., & Hutchins, G. (2000). A cross-linguistic PET study of tone perception. *Journal of Cognitive Neuroscience*, 12 (1), 207-222.

Gårding, E., Kratochvil, P., Svantesson, J.-O., & Zhang, J. (1986). Tone 4 and Tone 3

Discrimination in Modern Standard Chinese. *Language Speech*, 29, 281–293.

Gilber, R. C., & Liu, C. (2013). Thresholds of tone pitch contour discrimination for English listeners. *Proceedings of Meetings on Acoustics,* Volume 19. ICA 2013 Montreal Montreal, Canada.

Gottfried, T. L., & Suiter, T. L. (1997). Effects of linguistic experience on the identification of Mandarin Chinese vowels and tones. *Journal of Phonetics*, 25, 207–231.

Guo, L., & Tao, L. (2008). Tone production in Mandarin Chinese by American Students: A Case Study. *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)*, Volume 1, Pages 123-138.

Hallé, P., Chang, Y.-C., & Best C. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics,* 32 (2004) 395–421.

Hao, Y.-C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, 269-279.

He, Y. (2010). *Perception and Production of Isolated and Coarticulated Mandarin Tones by American Learners.* Thesis (Ph.D.), University of Florida.

He, Y., & Wayland, R. (2010). The Production of Mandarin Coarticulated Tones by Inexperienced and Experienced English Speakers of Mandarin. *Speech Prosody*. Chicago, IL, USA. May 10-14, 2010

Hothorn, T., Bretz, F., & Westfall, P. (2014) Simultaneous Inference in General Parametric Models. [http://cran.r-project.org/web/packages/multcomp/multcomp.pdf]

Howie, J. M. (1976). *Journal of the Acoustical Studies of Mandarin Vowels and Tones* (Cambridge University Press, Cambridge, MA).

Huang, T. (2004). *Language-specificity in auditory perception of Chinese tones.* Unpublished

Doctoral Dissertation, Ohio State University, Columbus.

Huang, J., & Holt, L. L. (2008). General perceptual contributions to lexical tone normalization. *Journal of the Acoustical Society of America,* 125 (6), Pages: 3983–3994.

Huang, T., & Johnson, K. (2010). Language Specificity in Speech Perception: Perception of Mandarin Tones by Native and Nonnative Listeners. *Phonetica,* 2010;67:243–267.

Jongman, A., Wang, Y., Moore, C. B., & Sereno, J. (2011). Perception and Production of Mandarin Chinese Tones. Submitted to: *Handbook of Chinese Psycholinguistics,* p209-217. Cambridge University Press.

Kiriloff, C. (1969). On the auditory discrimination of tones in Mandarin. *Phonetica,* 20, 63–67.

Krenmayr, A., Qi, B., Liu, B., Liu, H., Chen, X., Han, D., Schatzer, R., & Zierhofer, C. M. (2011). Development of a Mandarin tone identification test: Sensitivity index d' as a performance measure for individual tones. *International Journal of Audiology,* 50: 155–163.

Kuhl, P., & Iverson, P. (1995). Linguistic experience and the perceptual magnet effect, In W. Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (121-154). Baltimore: York.

Kuo, Y.-C., Rosen, S., & Faulkner, A. (2008). Acoustic cues to tonal contrasts in Mandarin: Implications for cochlear implants. *Journal of the Acoustical Society of America,* 123 (5).

Kuznetsova, A., Brockhoff, B. P., & Christensen, R. H. B. (2014). lmerTest: Tests for random and fixed effects for linear mixed effect models [http://cran.r-project.org/web/packages/lmerTest/index.html]

Lai, Y., & Zhang, J. (2008). Mandarin Lexical Tone Recognition: The Gating Paradigm. *Kansas Working Papers in Linguistics*, Vol. 30, p. 183.

Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather and A. James (eds.): N*ew Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second Language Speech*. Amsterdam: University of Amsterdam, 305-341.

Lee, C.-Y. (2001). *Lexical Tone in Spoken Word Recognition: A View from Mandarin Chinese.* Providence: Brown University Dissertaion.

Lee, C.-Y., Tao, L., & Bond, Z. S. (2006). Perception of acoustically modified Mandarin tones. *Journal of the Acoustical Society of America,* 119, 3243.

Lee, C.-Y., Tao, L., & Bond, Z. S. (2008). Identification of acoustically modifie Mandarin tones by native listeners. *Journal of Phonetics,* 36, 537-563.

Lee, C.-Y., Tao, L., & Bond, Z. S. (2010 a). Identification of acoustically modified Mandarin tones by non-native listeners. *Language Speech*, 53(Pt 2):217-43.

Lee, C.-Y., Tao, L., & Bond, Z. S. (2010 b). Identification of multi-speaker Mandarin tones in noise by native and non-native listeners. *Speech Communication* 52, 900–910.

Li, C.N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-Speaking children. *Journal of Child Language*,4, 185-199.

Lin, M. C., Yan, J. Z., & Sun, G. H. (1984). A Primary experiment on the stress pattern of normal disyllabic words in Mandarin. *Dialect*, No.1, pp 57-73.

Lin, T., & Wang, Y.-S. (1985). Shengdiao ganzhi wenti tone perception. *Zhongguo Yuyan Xuebao*, **2**, 59–69.

Littell, R., Stroup, W., Freund, R. (2002). *SAS for Linear Models, Fourth Edition.* SAS Publishing.[http://faculty.ucr.edu/~hanneman/linear_models/c4.html]

Liu, C. (2013). Just noticeable difference of tone pitch contour change for English- and Chinese-native listeners. *Journal of the Acoustical Society of America,* 134, 3011.

Liu, H.-M., Tsao, F.-M., & Kuhl, P. K. (2007). Acoustic Analysis of Lexical Tone in Mandarin Infant-Directed Speech. *Developmental Psychology, Vol. 43, No. 4, 912–917.*

Liu, H.-M., Chen, Y., & Tsao, F.-M. (2014). Developmental Changes in Mismatch Responses to Mandarin Consonants and Lexical Tones from Early to Middle Childhood. *PLoS ONE,* 9(4): e95587. doi:10.1371/journal.pone.009558.

Liu, L., Peng, D., Ding, G., Jin, Z., Zhang, L., Li, K., & Chen, C. (2006). Dissociation in the neural basis underlying Chinese tone and vowel production. *NeuroImage*. Volume 29, Issue 2, Pages 515–523.

Lu, J.-M. (1992). *On the Perception and Production of Mandarin Tones by Adult English-speaking Learners.* Michigan State University.

Macmillan, N. A. & Creelman, C. D. (2005). *Detection Theory: A User's Guide.* Psychology Press; 2nd edition.

Massaro, D. W., Cohen, M. M., & Tseng, C.-Y. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics*, 13, 267–289.

Mattock, K., & Burnham, D. (2006). Chinese and English Infants' Tone Perception: Evidence for Perceptual Reorganization*.* *Infancy*, 10 (3). 241-265. ISSN 1532-7078.

McGinnis, S. (1996). Tonal Distinction Errors by Beginning Chinese Language Students: A Comparative Study of American English and Japanese Native Speakers. *Chinese pedagogy: An emerging field*, Columbus, Ohio: Foreign Language Publications. 81-91.

McGuire, G. (2010). A Brief Primer on Experimental Designs for Speech Perception Research*. UCSC Linguistics Research Center.*

Miracle, W. C. (1989). Tone production of American students of Chinese: A preliminary acoustic study. *Journal of the Chinese Language Teachers Association*, 24, 49–65.

Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *Journal of the Acoustical Society of America*, 102, 1864–1877.

Ning, L.-H., Shih, C., & Loucks, T. M. (2014). Mandarin Tone Learning in L2 Adults: A Test of Perceptual and Sensorimotor Contributions. *Speech Communication*, Volumes 63 – 64, Pages 55 – 69.

Oller, D. K. & Smith, L. (1977). Effect of final-syllable position on vowel duration in infant babbling. *Journal of the Acoustical Society of America.* Volume 62, 994-997.

Peng, S.-H. (1996). *Phonetic Implementation and Perception of Place Coarticulation and Tone Sandhi.* PhD Dissertation. Ohio State University.

Peng, S.-H. (2000). Lexical versus `phonological' representations of Mandarin sandhi tones. In Papers in Laboratory Phonology {V}: *Acquisition and the Lexicon*, 152-167.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Shen, J., Deutsch, D., & Rayner, K. (2013). On-line perception of Mandarin Tones 2 and 3: Evidence from eye movements. *Journal of the Acoustical Society of America,* 133 (5), Pages: 3016–3029.

Shen, X. S. (1989). Toward a register approach in teaching Mandarin tones, *Journal of the Chinese Language Teachers Association,* 24, 27–47.

Shen , X . S. (1990). Tonal coarticulation in Mandarin. *Journal of Phonetics ,* 18 , 281 – 295

Shen, X. S., & Lin, M. C. (1991). A perceptual study of Mandarin tones 2 and 3. *Lang Speech* 34, 145–156.

Shih, C. (1987). The phonetics of Chinese tonal system. *Bell Laboratories, Tech. Memo*.

Shih, C. (1988). Tone and Intonation in Mandarin. In N. Clements ed. *Working Papers of the Cornell Phonetics Laboratory 3*. 83-109.

Shih, C., & Sproat, R. (1992). Variations of the Mandarin Rising Tone. *Proceedings of the IRCS Workshop on Prosody in Natural Speech*, 92(37), 193-200, University of Pennsylvania.

So, C. K., & Best, C. T. (2010). Cross-language Perception of Non-native Tonal Contrasts: Effects of Native Phonological and Phonetic Influences. *Language and Speech*, 53(2), 273–293.

So, C. K., & Best, C. T. (2011). Categorizing Mandarin Tones into Listeners' Native Prosodic Categories: The Role of Phonetic Properties. *Poznań Studies in Contemporary Linguistics,* 47(1), 2011, 133–145.

So, C. K., & Best, C. T. (2014). Phonetic Influences on English and French Listeners' Assimilation of Mandarin Tones to Native Prosodic Categories. *Studies in Second Language Acquisition*, Vol. 36 Issue 2, p195-221, 27p.

Stagray, J. R., & Downs, D. (1993). Differential sensitivity for frequency among speakers of a tone and a nontone language. *Journal of Chinese Linguistics,* 21, 143-163.

Tseng, S.-C., & Lee, T.-L. (2013). Tones of Recued T1-T4 Mandarin Disyllables. *International Journal of Computational Linguistics and Chinese Language Processing*, 18:3. 081-105.

Trofimovich, P., Baker, W., Flege, J. E., & Mack, M. (2003). Second-language sound learning in children and adults: learning sounds, words, or both? In B. Beachley, A. Brown, &F. Colin (Eds.), Proceedings of the 27th Boston University Conference on Language Development

(775-786). Someville, MA: Cascadilla Press

Wang, W. S.-Y., & Li. K.-P. (1967). Tone 3 in Pekinese. *Journal of Speech and Hearing Research*, 10:629-636.

Wang, Y. J. (2004). Tone Pattern and Word Stress in Mandarin. International Symposium on Tonal Aspects of languages: With Emphasis on Tone Languages. Beijing, China.

Wang, Y., Behne, D. M., Jongman, A., & Sereno, J. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics*, Volume 25, Issue 03, pp 449-466.

Wang, Y., Jongman, A., & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language*, 78, 332-348.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America,* Volume 113, Issue 2, 1033-1043;

Wang, Y., Jongman, A., & Sereno, J. (2006). Second language acquisition and processing of Mandarin tone. In E. Bates, L. Tan, and O. Tzeng (Eds.), *Handbook of Chinese Psycholinguistics*, Cambridge: Cambridge University Press, 250-257.

Wang, Y., Sereno, J. A., Jongman, A., & Hirsch, J. (2003). fMRI Evidence for Cortical Modification during Learning of Mandarin Lexical Tone. *Journal of Cognitive Neuroscience,* Volume 15 Issue 7, 1019-1027.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America,*. 106, 3649–3658.

Whalen, D. H., & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, 49, 25-47.

White, C. M. (1981). Tonal perception errors and interference from English intonation. *Journal of the Chinese Language Teachers Association*, 16, 27–56.

Wong, P. (2008). *Development of lexical tone production in disyllabic words by 2- to 6-year-old Mandarin-speaking children* (Doctoral dissertation). The Graduate Center of the City University of New York.

Wong, P. (2012a). Monosyllabic Mandarin Tone Productions by 3-Year-Olds Growing Up in Taiwan and in the United States: Interjudge Reliability and Perceptual Results. *Journal of Speech, Language, and Hearing Research*, Vol. 55, 1423–1437.

Wong, P. (2012b). Acoustic characteristics of three-year-olds' correct and incorrect monosyllabic Mandarin lexical tone productions. *Journal of Phonetics*, 40, 141–151.

Wong, P. C. M., Parsons, L. M., Martinez, M., & Diehl, R. L. (2004). The role of the insular cortex in pitch pattern perception: The effect of linguistic contexts. *The Journal of Neuroscience*, 24, 9153-9160.

Wong, P. C. M., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics,* 28 (2007), 565–585.

Wong, P., Schwartz, R. G., & Jenkins, J. J. (2005). Perception and production of lexical tones by 3-year-old, Mandarin-speaking children. *Journal of Speech, Language, and Hearing Research,* 48, 1065-1079.

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical Perception of Lexical Tones in Chinese Revealed by Mismatch Negativity. *Neuroscience,* 170 (2010) 223–231.

Xu, Y. (1993). *Contextual Tonal Variations in Mandarin Chinese*. PhD Dissertation, University of Connecticut, 1993.

Xu, Y. (1994). Production and perception of coarticulated tones. *Journal of the Acoustical*

Society of America, Vol.95.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25: 61-83.

Xu, Y. (2004). Understanding Tone from the Perspective of Production and Perception. *Language and Linguistics,* 5.4:757-797.

Xu, Y. (2006). Tone in connected discourse. In *Encyclopedia of Language and Linguistics, 2nd Ed.* K. Brown. Oxford: Elsevier. 12: 742-750.

Xu, Y., & Sun, X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America,* 111:1399-1413.

Xu, Y., & Wang, Q. E. (2001).Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication,* 33:319-337.

Xu, Y., & Wang, Q. E. (2003). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication,* 33 (2001) 319-37.

Xu, Y. S., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America,*120 (2).

Yang, B. (2010). *A model of Mandarin tone categories--a study of perception and production.* PhD dissertation, University of Iowa, 2010.

Yue-Hashimoto, A. O. (1986). Shengdiao jiaofa de Shangque. [A discussion of the teaching of tone]. Diyijie Guoji Hanyu Jiaoxue Taolunhui Lunwenxuan,[In the first international symposium on teaching Chinese as a foreign language, Beijing: Beijing Language Institute], 229-235.

Yip, M. (2002). *Tone*. Cambridge Univ. Press.

Appendices

Appendix I

Experimental Tasks Reading Lists

Please read each following character two times with a 2 second interval between each utterance:

Simplified Character	Syllable	Pinyin	Gloss
妈	Tone 1 /ma/	mā	mother
麻	Tone 2 /ma/	má	hemp
马	Tone 3 /ma/	mǎ	horse
骂	Tone 4 /ma/	mà	scold
压	Tone 1 /ja/	yā	press
牙	Tone 2 /ja/	yá	tooth
雅	Tone 3 /ja/	yǎ	elegant
讶	Tone 4 /ja/	yà	surprised

Please read the following non words as naturally as possible, i.e. with no pause between the two syllables. Please read each non word two times.

Syllable	Pinyin
Tone 1 Tone 1 /ma.ja/	mā. yā
Tone 1 Tone 2 /ma.ja/	mā. yá
Tone 1 Tone 3 /ma.ja/	mā. yǎ
Tone 1 Tone 4 /ma.ja/	mā. yà
Tone 2 Tone 1 /ma.ja/	má. yā
Tone 2 Tone 2 /ma.ja/	má. yá
Tone 2 Tone 3 /ma.ja/	má. yǎ
Tone 2 Tone 4 /ma.ja/	má. yà
Tone 3 Tone1 /ma.ja/	mǎ. yā
Tone 3 Tone2	mǎ. yá

/ma.ja/	
Tone 3 Tone3 /ma.ja/	mǎ. yǎ
Tone 3 Tone4 /ma.ja/	mǎ. yà
Tone 4 Tone 1 /ma.ja/	mà. yā
Tone 4 Tone 2 /ma.ja/	mà. yá
Tone 4 Tone 3 /ma.ja/	mà. yǎ
Tone 4 Tone 4 /ma.ja/	mà. yà
Tone 1 Tone 1 /ja.ma/	yā. mā
Tone 1 Tone 2 /ja.ma/	yā. má
Tone 1 Tone 3 /ja.ma/	yā. mǎ
Tone 1 Tone 4 /ja.ma/	yā. mà
Tone2 Tone 1 /ja.ma/	yá. mā

Tone 2 Tone 2 /ja.ma/	yá. má
Tone 2 Tone 3 /ja.ma/	yá. mǎ
Tone 2 Tone 4 /ja.ma/	yá. mà
Tone 3 Tone 1 /ja.ma/	yǎ. mā
Tone 3 Tone2 /ja.ma/	yǎ. má
Tone 3 Tone3 /ja.ma/	yǎ. mǎ
Tone 3 Tone4 /ja.ma/	yǎ. mà
Tone 4 Tone 1 /ja.ma/	yà. mā
Tone 4 Tone 2 /ja.ma/	yà. má
Tone 4 Tone 3 /ja.ma/	yà. mǎ
Tone 4 Tone 4 /ja.ma/	yà. mà

Appendix II

Familiarity Task Reading Lists

Please read the following each character two times with 2 seconds interval between each utterances:

Simplified Character	Syllable	Pinyin	Gloss
依	Tone 1 /ji/	yī	according to
姨	Tone 2 / ji /	yí	aunt
椅	Tone 3 / ji /	yǐ	chair
易	Tone 4 / ji /	yì	easy
低	Tone 1 /ti/	dī	low
迪	Tone 2 / ti /	dí	di (sound borrowing)
底	Tone 3 / ti /	dǐ	bottom
地	Tone 4 / ti /	dì	ground

Please read the following non words as naturally as possible, i.e. with no pause between the two syllables. Please read each non word two times.

Syllable	Pinyin
Tone 1 Tone 1 /yi.di/	yī. dī
Tone 1 Tone 2 /yi.di/	yī. dí
Tone 1 Tone 3 /yi.di/	yī. dǐ
Tone 1 Tone 4 /yi.di/	yī. dì
Tone 2 Tone 1 /yi.di/	yí. dī
Tone 2 Tone 2 /yi.di/	yí. dí
Tone 2 Tone 3 /yi.di/	yí. dǐ
Tone 2 Tone 4 /yi.di/	yí. dì
Tone 3 Tone1 /yi.di/	yǐ. dī
Tone 3 Tone2	yǐ. dí

/yi.di/	
Tone 3 Tone3 /yi.di/	yǐ. dǐ
Tone 3 Tone4 /yi.di/	yǐ. dì
Tone 4 Tone 1 /yi.di/	yì. dī
Tone 4 Tone 2 /yi.di/	yì. dí
Tone 4 Tone 3 /yi.di/	yì. dǐ
Tone 4 Tone 4 /yi.di/	yì. dì

Appendix III

The instruction of Task 1 (Tones in the monosyllable)

Thank you for agreeing to participate in this experiment. Whenever you get a message screen like this, press SPACEBAR to move to the next screen. Press SPACEBAR now. This experiment is composed of many trials. In each trial, you'll hear one Mandarin syllable. Your task is to decide which Mandarin tone is being said. Please response as soon as possible. In other words, you need to decide what the Mandarin tone is for each syllable that you will hear. If it sounds like the **Tone 1**; you'll hit the **D** key on the keyboard. If it sounds like the **Tone 2**, you'll hit the **F** key on the keyboard. If it sounds like the **Tone 3**, you'll hit the **J** key on the keyboard. If it sounds like the **Tone 4**, you'll hit the **K** key on the keyboard. "¯ (Tone 1)", " / (Tone 2)", " v (Tone 3)", and " \ (Tone 4)", four tone contours will be displayed on the screen as a visual aid to help you. For example: you might hear **mā** ", then you'll hit the **D** key, which is marked with "ā". Likewise, you might hear **yá**, then you'll hit the **F** key, which is marked with "á". Likewise, you might hear **mǎ**, then you'll hit the **J** key which is marked with "ǎ". Likewise, you might hear yà, and then you'll hit the K key, which is marked with "à". After the syllable finishes playing, you will only have two seconds to give your response before the next trial begins. This means you won't have time to think carefully about your answer. Just respond based on your first impulse. Please respond as soon as possible. You may even respond while the syllable is still playing. Press SPACEBAR to begin the experiment.

The instruction of Task 2 (Tones in the initial syllable)

Thank you for agreeing to participate in this experiment. Whenever you get a message screen like this, press SPACEBAR to move to the next screen. Press SPACEBAR now. This experiment is composed of many trials. In each trial, you'll hear a disyllabic (two syllables) nonsense word. Your task is to decide the tone of the first syllable. Please response as soon as possible. In other words, you need to decide what the Mandarin tone for the first syllable you have heard is. If it sounds like the **Tone 1**; you'll hit the **D** key on the keyboard. If it sounds like the **Tone 2**, you'll hit the **F** key on the keyboard. If it sounds like the **Tone 3**, you'll hit the **J** key on the keyboard. If it sounds like the **Tone 4**, you'll hit the **K** key on the keyboard. "¯ (Tone 1)", " / (Tone 2)", " v (Tone 3)", and " \ (Tone 4)", four tone contours will be displayed on the screen as a visual aid to help you. For example: you might hear **mā yá** , then you'll hit the **D** key, which is marked with "ā". Likewise, you might hear **yá mā**, then you'll hit the **F** key, which is marked with "á". Likewise, you might hear **mǎ yá**, then you'll hit the **J** key which is marked with "ǎ". Likewise, you might hear **yà mǎ**, and then you'll hit the K key, which is marked with "à". After the syllable finishes playing, you will only have two seconds to give your response before the next trial begins. This means you won't have time to think carefully about your answer. Just respond based on your first impulse. Please respond as soon as possible. You may even respond while the syllable is still playing. Press SPACEBAR to begin the experiment.

The instruction of Task 3 (Tones in the final syllable)

Thank you for agreeing to participate in this experiment. Whenever you get a message screen like this, press SPACEBAR to move to the next screen. Press SPACEBAR now. This experiment is composed of many trials. In each trial, you'll hear a disyllabic (two syllables) nonsense word. Your task is to decide the tone of the second syllable. Please response as soon as possible. In other words, you need to decide what the Mandarin tone for the second syllable you have heard is. If it sounds like the **Tone 1**; you'll hit the **D** key on the keyboard. If it sounds like the **Tone 2**, you'll hit the **F** key on the keyboard. If it sounds like the **Tone 3**, you'll hit the **J** key on the keyboard. If it sounds like the **Tone 4**, you'll hit the **K** key on the keyboard. "¯ (Tone 1)", " / (Tone 2)", " v (Tone 3)", and " \ (Tone 4)", four tone contours will be displayed on the screen as a visual aid to help you. For example: you might hear **yá mā**, then you'll hit the **D** key, which is marked with "ā". Likewise, you might hear **mā yá**, then you'll hit the **F** key, which is marked with "á". Likewise, you might hear **yá mǎ**, then you'll hit the **J** key which is marked with "ǎ". Likewise, you might hear **mǎ yà**, and then you'll hit the K key, which is marked with "à". After the syllable finishes playing, you will only have two seconds to give your response before the next trial begins. This means you won't have time to think carefully about your answer. Just respond based on your first impulse. Please respond as soon as possible. You may even respond while the syllable is still playing. Press SPACEBAR to begin the experiment.

Appendix IV

Linguistic Background Questionnaire

1. Subject code: __________

2. Date: ________ Major: ________ Gender: male female Year of Birth: 19____

3. Which country did you live in before the age 7? __________

4. Do you have previous experience living in a Chinese-speaking country? Yes. No.

If Yes, for how long? ____________________

5. Which languages have you learned so far? (Including Chinese)

Languages	Age of learning	Country	Instruction location	From who (parents, teacher, etc.)	Overall Proficiency 1-unfluent, 7-fluent
Chinese					1 2 3 4 5 6 7
					1 2 3 4 5 6 7
					1 2 3 4 5 6 7

6. Are you a professional musician? Yes. No.

7. What is the biggest difficulty for you regarding learning Mandarin pronunciation? Why?

____

8. Do you have any suggestions in terms of Chinese drills or formal instruction concerning improving your pronunciation?

____

9. Are there any additional difficulties you have encountered while studying Chinese?

____

LIU, MENG

E-mail: mengliu2@illinois.edu

EDUCATION

PhD in East Asian Languages and Cultures — Aug. 2012--Present
University of Illinois at Urbana-Champaign (UIUC)

MA in East Asian Languages and Cultures — Sept.2010--Jul. 2012
Indiana University—Bloomington (IUB)

BA in Teaching Chinese as a Second Language — Sept. 2006--Jul. 2010
Beijing Language and Culture University (BLCU)

STANDARDIZED TESTS:

TOEFL: 104 (Reading: 25, Listening: 27, Speaking: 24, Writing: 28) — Date: 12/27/2011
TEPAIC: Certification level 2 – Satisfactory Certification — Date: 08/25/2010
Chinese (Mandarin) test: A level (92 .9/100) — Date: 05/10/2010
GRE: 1260 (Verbal: 550 75%, Quantitative: 710 73%, Analytical Writing: 4.0 41%) — Date: 10/24/2009

WORK EXPERIENCE

- Chinese Lecturer for CHIN 306 (business Chinese) at UIUC — Jan.2015-Present
- Chinese Lecturer for CHIN 305/490 (business Chinese) at UIUC — Aug.2013-May 2014
- Instructor of Business Chinese for EMBA — Apr. 2 &16, 2014
- Lead Instructor for FCI C301/302(third year Chinese) at IUB — May 2013- Jul.2013
- Chinese Lecturer for CHIN305/CHIN306 (third year Chinese) at UIUC — Aug.2012-May 2013
- Instructor of Business Chinese for EMBA — Mar. 1, 2013
- Lead Instructor for FCI C101/102 (first year Chinese) at IUB — Jun.2012- Aug.2012
- Associate Instructor for C201/202(second year Chinese) at IUB — Aug. 2011- May 2012
- Language Facilitator at Global Village Living-Learning Center, IUB Foster Residence Center — Oct. 2010- May 2012
- Associate Instructor for FCI C301/302 (third year Chinese)at IUB — Jun.2011- Aug.2011
- Associate Instructor for C101/102 (first year Chinese)at IUB — Aug. 2010- May 2011
- Instructor of China 101 (IU SPEA) — Feb. 2011
  Introduced Chinese culture, society, history and economics for the manager of IU media relations, as well as IU School of Journalism faculty.
- Tutor for international students from the US, Canada, Britain, Germany, Spain, Switzerland, Japan, and Korea
  Taught Chinese and helped them prepare for HSK Test — Apr. 2007- Jul. 2010
- Beijing--Harvard Chinese Institution
  Teaching Assistant for the 4th level and intermediate learners — Jul. 2008, Jun. 2009
- Chinese Language Training Institute, BLCU — May 2009
  Interned as a teacher of Listening and Speaking curriculum for intermediate level classes
- Tutor for Korean OTO Company — Dec.2008
  Taught Business Chinese to employees

LIU, MENG
E-mail: mengliu2@illinois.edu

CONFERENCE PRESENTATION

- Shih, C.L., Wu, D., **Liu, M.**, Yang Y.H., Roy, J. (2014). Mapping Second Language Accents. Focal Point Bilingualism Symposium – I. University of Illinois at Urbana-Champaign, Illinois, May 3, 2014.
- **Liu, M.** (2014). Practical Approaches to Teaching Business Chinese by Using Technology. The 16th Annual CIBER Business Language Conference. University of Utah, Park City, Utah, April 25, 2014.
- **Liu, M.** (2014). *Prosodic Factors in Foreign Accents of L1 English learners of Chinese.* The Illinois Language and Linguistics Society's 6th Conference. University of Illinois at Urbana-Champaign, Illinois, April 6, 2014.
- **Liu, M.** (2013). *The perception of Mandarin Tones by American learners of Chinese.* American Council on the Teaching of Foreign Languages (ACTFL). Orlando, Florida, November 22, 2013.
- **Liu, M.** (2013). *The perception of Mandarin tones in contexts.* The Illinois Language and Linguistics Society's 5th Conference. University of Illinois at Urbana-Champaign, Illinois, April 5, 2013.
- **Liu, M.** (2012). *The perception of Mandarin tones by L1 English learners of Chinese.* The 24th North American Conference on Chinese Linguistics. The University of San Francisco, California. June 8, 2012.

RESEARCH EXPERIENCE

- Participated in the Language & Cognition Lab meeting, IUB — Sept. 2011-May 2012
- *Undergraduate Guiding Norm of Higher Education in Chinese Major Chinese*, guided by Prof. Shi Jiawei, BLCU. — Sept. 2009
  Collected statistics and data and utilized quantitative and qualitative methods to analyze curriculum design of TCSL

PROFESSIONAL ORGANIZATION MEMBERSHIP

- American Council on The Teaching of Foreign Languages (ACTFL) — Aug. 2013-Present
- Chinese Language Teachers Association (CLTA) — Jan.2013-Present

EXTRACURRICULAR ACTIVITIES

LEADERSHIP EXPERIENCE

- Track and Field Team, BLCU: Team leader — Sept.2008- Jul. 2010
  Competed on behalf of BLCU at the 45th, 46th and 47th Beijing Intramural Competition
- Student Union of the College of Humanities and Social Sciences, BLCU — Sept.2007-Sept.2008
  President of the Study Division
  Organized the Annual University Debate, Freshman Advising Meeting, and Spoken English Contest

VOLUNTEER EXPERIENCE

- Chinese language immersion instructor at Illinois state Global Fest 2014 — Mar. 1, 2014
  Introduced Chinese geography, food, language and culture for Illinois state middle to high school students

---

- The Illinois Language and Linguistics Society's 5th Conference Volunteer — Apr. 5-7, 2013
  Helped with the registration procedure, logistics, and guided presenters
- English to Mandarin Translation — Sept. 2012
  Translated series of materials for adoptive parents of a Chinese girl
- Indiana University—Bloomington East Asian Studies Center Volunteer — Feb. 2012
  Presented Chinese cultures (New Year traditions, legends, zodiacs, etc.) to elementary school students
- 2008 Beijing Paralympics Volunteer — Aug. 2008-Sept. 2008
  Assisted the Executive Manager of the Department for Foreign Athlete Affairs
- 2008 Beijing Olympics Volunteer — Jul. 2008-Aug. 2008
  Interpreted and assisted full-time for the Canadian Olympic Delegation
- "Good Luck Beijing" Sports Meeting Volunteer — Nov. 2007-Dec. 2007
  Interpreted and assisted full-time for the Swiss and Australian Gymnastics Delegation
- Member in the Olympic Volunteer Association of BCLU — Sept. 2006-Oct. 2007
  Participated in teaching English in the Ba Lizhuang Community, Chao Yang District, Beijing
  Taught Business English to servers at Yan Shan Hotel, Beijing

HONORS and AWARDS

ACADEMIC AWARDS

- Ranked as "Excellent Teacher" by students at UIUC. — Spring 2014
  http://cte.illinois.edu/teacheval/ices/pdf/Spring14List.pdf
- CIBER Doctoral Student Travel Award ($850) — Feb. 25, 2014
- EALC Departmental Conference Travel Award ($400) — Nov. 1, 2013
- Ranked as "Excellent Teacher" by students at UIUC. — Fall 2013
  http://cte.illinois.edu/teacheval/ices/pdf/Fall13List.pdf
- Ranked as "Excellent Teacher" by students at UIUC. — Spring 2013
  http://cte.illinois.edu/teacheval/ices/pdf/Spring13List.pdf
- The finalists of CLTA Walton Presentation Prize — May 1, 2013
- CTE (Center for Teaching Excellence) Graduate Teacher Certificate — May 1, 2013
- EALC Summer Research Fellowship, UIUC ($1500) — Apr. 5, 2013
- Ralph Tyler Award for "Best Chinese Language Teaching Assistant" of UIUC ($500) — Mar. 25, 2013
- Fellowship ($3000) East Asian Languages and Cultures, UIUC — Aug.2012
- "Excellent Presenter" of Chinese Musical Instruments at the IU World Language Festival — Apr.2011
- "Excellent Student" of the College of Humanities and Social Sciences, BLCU (top 10%) — Oct. 2008, Nov. 2009
- 2nd Prize of Comprehensive Annual Scholarship, BLCU (top 10%) — Sept.2007, Sept.2008, Sept.2009
- Zhen Hai Annual Scholarship for Excellent Students, BLCU (20 people out of the whole University) — Sept. 2007, Sept. 2008, Sept. 2009
- Honorable Mention in "Challenge Cup" extra-curricular academic competition, BLCU — May 2009
- Member of the Beijing Excellent Youth League (top 3% Youth Leagues in Beijing) — May 2008

LIU, MENG

E-mail: mengliu2@illinois.edu

ATHLETIC AWARDS

- 5th place in the 4×400m relays and 7th place in the 4×100m relays at the 47th Beijing Intramural Competition — May 2009
- 1st, 2nd, & 3rd place in the 800m races, 4×100m relays and 400m races (respectively) in the 28th, 29th, 30th BLCU Annual Sports Meeting — Apr. 2007, Apr. 2008, Apr. 2009
- 2nd, 1st place in the 5000m Annual Winter Running event, BLCU — Dec. 2007, Dec. 2008

OTHER INFORMATION

- Research skills- Praat, R, Matlab, SPSS, SYSTAT
- CPU Skills – Proficient in Microsoft Office Suite: Word, Excel, Access, Power Point.
- Languages – Professional English (fluent), Spanish, German and Japanese (basic)
- Arts – 10th level certificate in Keyboard (music) from the Central Conservatory, China.

REFERENCE

- Chilin Shih
  Associate Professor of East Asian Languages and Cultures, UIUC
  EALC 2090 Room A, FLB 707 South Mathews Avenue, Urbana, IL 61801
  cls@illinois.edu

- Jennifer Li-Chia Liu
  Director of the Chinese Language Program, EALC, Harvard University
  5 Bryant St. #207. Cambridge, MA 02138
  liu02@fas.harvard.edu

- Vivian Ling
  Interim Director, Chinese Flagship Program
  1011 E Third St, Bloomington, IN 47405
  vling@indiana.edu